DOCUMENT RESUME

ED 148 343 IR 005 444

TITLE Computer Science and Technology: A Survey of Eleven Government-Developed Data Element Dictionary/Directory Systems.
INSTITUTION National Bureau of Standards (DOC), Washington, D.C. Inst. for Computer Sciences and Technology.
REPORT NO NBS-SP-500-16
PUB DATE Aug 77
NOTE 119p.; Appendix A of limited reproducibility; Computer Science & Technology Series
AVAILABLE FROM Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (Stock no. C13.10:500-16)

EDRS PRICE MF-$0.83 HC-$6.01 Plus Postage.
DESCRIPTORS Agencies; Dictionaries; Directories; *Electronic Data Processing; Federal Government; Guidelines; Programing Languages; *State of the Art Reviews
IDENTIFIERS *Computer Software

ABSTRACT

This report presents the current state of the art of government developed Data Element Dictionary/Directory (DED/D) systems. DED/D's are software tools used for managing and controlling information and data. The introduction of the report includes a list of the government agency systems surveyed and a summary matrix presenting each system's general characteristics. A comparative list focusing on the technical characteristics of each system, a narrative description highlighting each system's special characteristics and the agency's experience with it, and scope notes and definitions of the terms referred to in the report are appended. Information presented in this report is intended to serve both the technical and administrative automated data processing community. (Author/JPF)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED148343

COMPUTER SCIENCE & TECHNOLOGY:

A Survey of Eleven Government-Developed Data Element Dictionary/Directory Systems

Institute for Computer Sciences and Technology
National Bureau of Standards
Washington, D.C. 20234

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

U.S. DEPARTMENT OF COMMERCE, Juanita M. Kreps, Secretary

Dr. Sidney Harman, Under Secretary
Jordan J. Baruch, Assistant Secretary for Science and Technology

NATIONAL BUREAU OF STANDARDS, Ernest Ambler, Acting Director

Issued August 1977

IR005-444

Reports on Computer Science and Technology

The National Bureau of Standards has a special responsibility within the Federal Government for computer science and technology activities. The programs of the NBS Institute for Computer Sciences and Technology are designed to provide ADP standards, guidelines, and technical advisory services to improve the effectiveness of computer utilization in the Federal sector, and to perform appropriate research and development efforts as foundation for such activities and programs. This publication series will report these NBS efforts to the Federal computer community as well as to interested specialists in the academic and private sectors. Those wishing to receive notices of publications in this series should complete and return the form at the end of this publication.

National Bureau of Standards Special Publication 500-16
Nat. Bur. Stand. (U.S.), Spec. Publ. 500-16, 111 pages (Aug. 1977)
CODEN: XNBSAV

Library of Congress Cataloging in Publication Data

Institute for Computer Sciences and Technology.

A survey of eleven government-developed data element dictionary/directory systems.

(Computer science & technology) (National Bureau of Standards special publication ; 500-16)

Supt. of Docs. no.: C13.10:500-16

1. Data base management—Computer programs. I. Title. II. Title: Data element dictionary/directory systems. III. Series: United States. National Bureau of Standards. Special publication ; 500-16.
QC100.U57 no. 500-16 [QA76.9.D3] 602'.1 [001.6'425] 77-608155

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1977

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402

Stock No. 003-003-01817-1

ACKNOWLEDGMENT

This report is the product of the efforts of the Federal Information Processing Standards Task Group 17 "Data Element Dictionary/Directories. Particular recognition is given to the following task group members and contributors:

Thomas Bergin, Veterans Administration
Paul Brown, U.S. Civil Service Commission
Thomas A. Chittenden, Department of Transportation
Sidney D. Clemons, Library of Congress
Warren Crow, State of New York Dept. of Motor Vehicles
Nicholas Deakos, Army Materiel Development & Readiness Command
M. Peter de Regt, National Oceanic and Atmospheric Administration
Carol Erdahl, Veterans Administration
Larry Hellel, Department of Agriculture
George Heller, Bureau of the Census
Aaron Hochman, Department of Defense
William H. Kenworthey, Department of Defense
Thomas M. Kurihara, Department of Agriculture
Belkis Leong-Hong, National Bureau of Standards
Beatrice Marron, National Bureau of Standards
Robert Mattes, Social Security Administration
Donald J. McCaffrey, Federal Power Commission Consultant
Hazel McEwen, National Bureau of Standards
John Milanese, Economic Development Administration
Carolyn D. Moore, Agency for International Development
William C. Morgan, U.S. Coast Guard
Walter Peacock, Federal Power Commission
Fernando Puente, Department of Defense
John Roberts, State of New York Dept. of Motor Vehicles
Frank Tagler, Department of Defense
Mary Thomas, Federal Deposit Insurance Corporation
Roxanne Williams, Department of Agriculture
John Yienger, Federal Power Commission
Edith Young, Commission on Federal Paperwork
Virginia Zylla, Veterans Administration

PREFACE

During the past two decades, advances in information processing and communication technologies have substantially increased the ability of organizations to collect, store, and process data in support of a vast and ever increasing number of management programs. Along with this increased ability has come recognition of the need for automated tools to manage data and information which in turn has led to a variety of software products. This study was undertaken to provide a state-of-the-art survey of data element dictionary/directory systems developed by government agencies.

This report identifies data element dictionary/directory systems by name and source and provides a description of their features. This, in no case, implies a recommendation or endorsement by the National Bureau of Standards nor should this presentation be construed as a certification that any of these systems provide the indicated capabilities. The information is presented as furnished by the participating agencies and has been reviewed by them for accuracy and clarity.

CONTENTS

1. Introduction.....1
2. The Data Element Dictionary/Directory as a Tool for Data Management..6
3. Survey of Government Data Element Dictionary/Directory Systems.....8
4. Summary.....11
FIGURE 1 - Summary Matrix.....5

Appendix A - Feature List
General Information.....14
System Features - Computer Hardware Environment.....15
System Features - Computer Software Environment.....16
System Features - User Environment.....17
Data Element Attributes Described by System.....18
Element Occurrences Listed.....19
Output/Products.....20

Appendix B - Scope Notes for the Feature List.....21

Appendix C - Narrative System Descriptions.....25
ADISDED - Associate Directorship for Information System Data Element Dictionary (U.S. Civil Service Commission).....26
ASCSDBD - Agriculture Stabilization and Conservation Service Data Base Directory (Department of Agriculture).....31
DARCOM DED/SYSCAD - U.S. Army Materiel Development and Readiness Command Data Element Dictionary/System Control and Documentation (Department of Defense - U.S. Army).....39
DEMS - Data Elements Management System (Department of State - Agency for International Development).....49
DLCS - Data Label Control System (Department of Transportation - U.S. Coast Guard).....55
DMV/DED - Department of Motor Vehicles Data Element Dictionary (State of New York Department of Motor Vehicles).....59

Appendix C - Narrative System Descriptions (continued)

FPCRIS - Federal Power Commission Regulatory Information System (Federal Power Commission).....65

LOGDRMS - Logistics Data Resource Management System (Department of Defense - Office of the Secretary of Defense).....69

STADES - Navy World-Wide Military Command and Control System (WWMCCS) Standard Data Element System (Department of Defense-U.S. Navy).....83

USDADBD - U.S. Department of Agriculture Data Base Directory (Department of Agriculture).....91

VADD - Veterans Administration Data Dictionary (Veterans Administration).....95

Appendix D - Scope Notes for the Narrative System Descriptions.....101

Appendix E - Definitions of Information Entities.....105

ABSTRACT

This report presents the current state-of-the-art of government-developed Data Element Dictionary/Directory (DED/D) systems. DED/D's are software tools used for managing and controlling information and data. Eleven DED/D systems are described, first using a side-by-side features presentation approach, and followed by narrative systems descriptions which highlight special capabilities and experiences with each system. Information presented in this report is intended to serve both the technical and administrative ADP community.

Keywords: ADP; automated data processing; computer software; data element dictionary; data element dictionary/directory; data management; software tool.

1. Introduction

An increasing awareness of the need to manage data resources has led to the development of systems that identify, describe, define, and relate the basic unit of information, the data element. To gain insight into these systems, Federal Information Processing Standards Task Group on Data Element Dictionary (FIPS TG 17) was given the task of developing guidelines for constructing data element dictionaries/directories (DED/D's). As a part of this effort, the task group is currently engaged in examining concepts, principles and applications of data resource management and data resource directories. Another of TG 17's assignment is to identify the relevant performance characteristics of the automated processes designed to use and maintain DED/D's.

The growing need for a tool to manage and control data resources has already affected Government managers, many of whom have developed and implemented DED/D's. Recognizing this, TG 17 established a sub-group to research, describe and compare as many of these Government DED/D systems as possible. Concurrently the Systems and Software Division, Institute for Computer Sciences and Technology, National Bureau of Standards (NBS) undertook to describe a representative set of commercially available DED/D's. The results of this effort are contained in NBS Special Publication 500-3, Technical Profile of Seven Data Element Dictionary/Directory Systems. Because of the obvious parallel between these two tasks, this document has been organized in a similar manner and draws upon material in NBS Special Publication 500-3.

Among tools for managing and controlling data resources, five general capabilities have been identified: data catalog, data dictionary, data directory, data dictionary/directory, and data resource directory. Since TG 17 wished to emphasize the use of DED/D's as tools for the management of data as an organizational resource - not as mechanical components of software systems - it excluded software modules that are essentially file predefinition tools for data base management systems or file management packages but are frequently called dictionaries. TG 17 realized that none of the systems examined was exclusively a catalog, dictionary, directory, or data resource directory, but a hybrid of these capabilities. Nevertheless, certain terms were found useful in portraying the orientation and scope of the DED/D within an agency's data resource management program. Concepts in data resource management are being explored in-depth by working groups within TG 17.

A significant number of the agencies responding to TG 17's initial request for participation had already implemented or were in the process of implementing DED/D's. TG 17 decided to use these DED/D Systems as the starting point of this survey. The systems included cover a wide range of purposes and capabilities, and reflect both differing philosophies of data management and differing levels of knowledge and experience in the field.

The purpose of this document is two fold: 1) to help the potential user in evaluating his requirements for a DED/D; and 2) to acquaint the reader with a sampling of existing systems and the associated features, as representative of the state-of-the-art within the Government community.

This is not an exhaustive list of government DED/D systems, but it reflects a selection of representative systems which TG-17 reviewed. Inclusion of a system does not imply recommendation or endorsement. Similarly, omission of a system does not imply that its capabilities are less than those of an included system.

ADISDED - Associate Directorship for Information Systems
Data Element Dictionary
U.S. Civil Service Commission

ASCSDBD - Agriculture Stabilization and Conservation
Service Data Base Directory
Department of Agriculture
Stabilization and Conservation Service

DARCOM DED/SYSCAD - Army Materiel Development and
Readiness Command Data Element
Dictionary and System Control
and Documentation
Department of Defense
Army Materiel Development and
Readiness Command (DARCOM)

DEMS - Data Element Management System
Department of State
Agency for International Development (AID)

DLCS - Data Label Control System
Department of Transportation
U.S. Coast Guard

DMB/DED - Department of Motor Vehicles Data Element Dictionary
State of New York
Department of Motor Vehicles

FPCRIS - Federal Power Commission Regulatory Information System
Federal Power Commission

LOGDRMS - Logistics Data Resource Management System
Department of Defense
Office of the Secretary of Defense (DoD LOGDESMO)

STADES - Navy World-wide Military Command and Control System
Standard Data Element System
Department of Defense
Navy Regional Data
(NARDAC Washington Naval Shipyard)

USDADBD - U.S. Department of Agriculture Data Base Directory
Department of Agriculture
Office of Automated Data Systems

VADD - Veterans Administration Data Dictionary
Veterans Administration

Documentation for all of the above systems is generally available with the exception of FPCRIS which is in the implementation phase.

The discussion on each system will be presented in two sections: 1) a feature list which focuses on the technical characteristics of each system appears as Appendix A; and 2) a narrative systems description which focuses on the development and implementation experiences of the implementor appears as Appendix C. Taken together, the feature list and the narrative system description represent a total DED/D system description. The feature list presents objectively the technical characteristics of the systems in a side-by-side format thereby lending itself to comparative analysis. The narrative describes the system in terms of the problems, needs, and resources of the particular agency for which the DED/D was implemented. In fact, it should be noted that each system narrative should be viewed only within the context. Further, it should be emphasized that the total system description is not intended to serve as a guide to system implementation or selection, but as an example of how common data management problems have been approached through systematic definition of data resources and their interrelationships.

This publication does not prescribe or recommend a level of implementation nor a given software system to any particular user. Instead, it suggests those factors which should be addressed by an organization interested in a DED/D system. Therefore, the reader should choose those factors which are relevant in developing his own functional requirements, design and performance specifications. It should be emphasized that it is important for the reader to first define his own position as closely as possible and then attempt a solution based on the experiences of those close to that position.

A summary matrix of the systems included in this survey is shown in Figure 1 and presents their general characteristics such as: operational mode, major hardware, degree of implementation and informational entities portrayed.

Appendix E provides applicable definitions of the terms referenced in this publication.

SUMMARY MATRIX

	ADIS DED	ASCS DBD	DARCOM	DEMS	DLCS	DMV DED	FPC RIS	LOGDRMS	STADES	USDA DBD	VADD
OPERATIONAL MODE: BATCH	YES	YES	YES	YES	YES	MANUAL SYSTEM	YES	YES	YES	YES	YES
INTERACTIVE	NO	NO	NO	NO	YES		YES	YES	YES	YES	NO
HARDWARE IMPLEMENTATION	UNIVAC SERIES 70	IBM 360	IBM 360	IBM 360	CDC 3300		IBM 370	IBM 360	HONEYWELL 6000	IBM 370	IBM 360
ENTITIES REFERENCED:											
PLANS/ PROGRAMS	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO
SYSTEMS	NO	NO	YES	YES	NO	YES	NO	YES	YES	NO	YES
APPLICATIONS	NO	NO	NO	NO	NO	NO	NO	YES	YES	NO	NO
PROCESSES (PROCEDURES)	NO	NO	YES	YES	NO	NO	YES	YES	YES	NO	YES
FILES	YES	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES
DOCUMENTS (TRANS-ACTIONS)	NO	NO	NO	NO	NO	NO	NO	YES	YES	NO	NO
FORMS	YES	YES	YES	YES	NO	NO	YES	YES	YES	NO	NO
REPORTS (FILE)	YES	NO	YES	YES	YES	NO	YES	YES	YES	NO	YES
RECORDS	NO	NO	NO	NO	NO	NO	NO	YES	YES	NO	NO
DATA ELEMENTS	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

FIGURE 1

2. The Data Element Dictionary/Directory as a Tool for Data Management

The recent introduction of the DED/D has provided management with an essential tool to inventory, document, and locate its data resources, thereby permitting the data to be managed as a resource of the organization.

Within this new class of tools, five general capabilities have been identified, that of:

1. data catalog - an organized listing, with or without a description, by full name of all data elements used by an organization;

2. data dictionary - an ordered collection of data element descriptions containing specific identification attributes.

3. data directory - an ordered collection of data element names, and/or identifiers, and attributes which provide location of the elements.

4. data dictionary/directory - an ordered collection of data elements that combines the features of a data catalog, data dictionary, and a data directory;

5. data resource directory - an ordered collection of informational entity identifiers and their attributes including those which provide location and interrelationship information. A data resource directory contains many of the features of a dictionary/directory, but is not limited to data elements insofar as informational entity content is concerned.

Since most of the eleven government-developed systems are of the data element dictionary/directory, focus will be placed on this group.

NBS Special Publication 500-3 has identified a set of characteristics for commercial DED/D's. The following subset of these characteristics generally applies to government DED/D's surveyed by TG 17. These systems:

1. contain a unique identification, a set of physical characteristics, and a textual description of each of the data elements;

2. show the relationships of elements to each other, and to other components of the system, e.g., programs, reports;

3. specify the source, location, usage and destination of the elements;

4. have validation and redundancy-checking capabilities;

5. contain security safeguards to control the accessibility to the data elements;

6. have reporting capabilities, such as:

   a. predefined management-oriented, statistical or summary reports,

   b. cross-reference reports,

   c. elements usage reports,

   d. audit trail reports,

   e. change-effect reports,

   f. error reports;

7. have retrieval capabilities, such as keywording, indexing, and online or batch querying.

These government systems also contain other features that distinguish each one from the other. For example, some systems have online interactive capabilities, and some are predominantly batch systems. The surveyed DED/D systems have as their main functions the control and management of data elements. Some of these systems are implemented without dependence on a data base management system (DBMS). However, there are others that are implemented as application systems using DBMS software. In both cases, the DED/D systems functions are not secondary to those of the DBMS.

NBS Special Publication 500-3 noted the following tangible benefits that can be derived from the use of a DED/D:

* Simple and effective control of the data elements;

* Reduction of data redundancy and inconsistency;

* Enforcement of standard usage;

* Enforcement of security safeguards and controlled accessibility of the data base;

* Determination of the impact on the total information activity from changes to data elements;

* Centralization of data elements as an aid in design and development of new systems.

* Consistency in documentation for data elements.

3. Survey of Government DED/D Systems

Most of the descriptions of the systems contained herein were provided by participants in the TG-17 effort.

A questionnaire was developed to collect the general characteristics of these systems. Based on the response from each interviewed participating agency representative, a set of criteria for inclusion was developed. The system had to:

1. be government-developed;

2. be implemented, or in the process of being implemented -- not just planned;

3. perform the functions of a DED/D - either at the element level, or at a higher level.

The information obtained about each system is presented in this report first in a side-by-side feature list format which consists mostly of short answers. Then it is presented in narrative form, which elaborates on unique aspects of the systems, and includes experiential elements of the development process.

It should be noted that the narrative systems descriptions are presented as they were originally submitted by the participating agencies with very minor stylistic or editorial changes. The content of these narrative descriptions may vary from system to system since agencies emphasize different aspects of the DED/D.

The scope-notes for the feature list are reproduced in Appendix B. It should be emphasized that the feature list represents an objective reporting of technical characteristics of each system, and it is not meant to convey critical or subjective system comparisons.

The feature list is organized under the following seven major headings:

1. General Information

   Information of a general nature is included in this section.

2. System Features - Hardware environment

   Information on the hardware environment in which the DED/D

is implemented is contained in this section.

3. System Features - Software environment

   Information about the software requirements for the DED/D is contained in this section.

4. System Features - User Environment

   Information on techniques and capabilities available to users is presented in this section.

5. Data Element Attributes Described by System

   Information about the characteristics of a data element as presented in a DED/D is contained in this section.

6. Element Occurrences Listed

   Additional information entities which are presented within the DED/D are listed.

7. Output/Products

   Typical reports or products that can be produced by the DED/D are listed in this section.

Appendix D contains the scope-notes for narrative systems descriptions. The narratives have been organized into the following sections:

Each of the DED/D systems described is different in scope and approach, thus the following categories could be used only as a guide for structuring descriptions of these systems.

1. Identification

   Information on the title and office of primary responsibility and on available documentation is included in this section.

2. Scope

   Reason for implementation, the background and history of the DED/D implementation, and the kind of implementation are included in this section.

3. ADP Resources (if needed)

   Hardware resources, including computer main frame, operating system, memory size, peripheral devices, special off line devices, software (whether specialized or nonspecialized), and system utilities are included in this section.

4. Content and Products

   Information on data relationships, organization of the DED/D, and standard reports or products are listed in this section.

5. Operational Considerations

   Factors relevant to data collection including source, method, edit, backup procedure, operating mode for update and query, and interface with other systems are described.

6. Security

   Restrictions on use, Freedom of Information, and Privacy Act considerations are detailed.

7. User Characteristics

   A description of the user organizations, and their functions, the data elements used and the reasons for using them.

8. Costs (this section is optional)

   If available, cost in dollars, and a description of personnel commitment necessary to implement and run the DED/D.

9. General Comments

   Comments not fitting the above categories which the proponents wish to make.

4. Summary

This publication has identified technical features of data element dictionary/directory systems (DED/D systems) and has presented side-by-side descriptions of 11 representative government-developed systems. In addition, narrative descriptions highlight special capabilities and experiences with these systems. Other TG 17 efforts are directed toward developing guidelines for future data resources management and data resource directives. This publication, however, is a presentation of the current status of government-developed DED/D systems. It is hoped that this practical approach will aid prospective users to relate their own situation to that of others who have evaluated and/or implemented systems, and to apply those factors which are relevant to their own set of circumstances.

Information contained in both the feature list and narrative descriptions is intended to serve a very wide range of users, including administrators, managers, analysts, programmers, specialists, and technicians.

It should be noted that although most agencies would be happy to provide available programs and documentation to other interested agencies -- either free of charge or for a nominal administrative charge -- these agencies are not prepared to provide maintenance for either "fixes" or enhancements. The acquiring agency will be "on its own," except for informal assistance. In addition, documentation may be sketchy or out of date.

It is also of interest to note that one system, the DMV/DED is not automated, but a set of manual procedures, which incorporate many of the features of the automated systems. Further ASCSDBD and LOGDRMS deal almost exclusively with documenting information entity data bases and are concerned with data elements only as components of these data bases.

The concept of viewing data as a resource has been growing in acceptance within the Federal Government. The number of participants in this study, and the interest shown in the work of FIPS TG 17 attest to the vital need for tools for managing and controlling data. The systems surveyed cover a wide range of capabilities and reflect differing philosophies of data management and differing levels of knowledge and experience in the field.

It should be reiterated that this publication does not prescribe or recommend a level of implementation or a given software system to any potential user. However, it does point to those factors which should be considered. It is up to the potential user to choose those factors which are relevant. By studying existing systems, it becomes possible to avoid pitfalls encountered by others, and to approach the problems of data management with potential solutions in hand.

APPENDIX A

FEATURE LIST

APPENDIX A
FEATURE LIST OF
DATA ELEMENT DICTIONARY/DIRECTORY
(DED/D) SYSTEMS
GENERAL INFORMATION

SYSTEM NAME	ADISDED	ASCSDBD	DARCOM	DEMS	DLCS	DMV/DED	FPCRIS	LOGDRMS	STADES	USDADBD	VADD
DEVELOPER	U.S. Navy Facilities Systems Office (FACSO)	U.S. Department of Agriculture Agriculture Stabilization and Conservation Service (ASCS) Data Systems Division	U.S. Army Material Development & Readiness Command (DoD)	Agency for International Development Office of Data Management, Information Management Division	U.S. Coast Guard Information Systems Division	NY State Dept. of Motor Veh. Data Standards and Controls Bureau Div. of Data Administration	Federal Power Commission Office of Regulatory Information Systems, and Planning Research Corp. (PRC)	Office of the Assistant Secretary of Defense (Installations and Logistics) DoD LOGDESMO	U.S. Navy Data Automation Command (NAVDAC) NARDAC Washington Naval Shipyard	U.S. Department of Agriculture Office of Automated Systems	Veterans Adm. Dept. of Data Management Privacy and Data Adm. Division
ADDRESS	Data Base Administrator Bureau of Manpower, Information Systems U.S. Civil Ser. Commission Wash., D.C.20415	Data Systems Division ASCS USDA Washington, D.C. 20250	Commander Automated Logistics Management Activity Attn:DRIAL-DED P.O. Box 1578 St. Louis, MO.	Information Management Div. Office of Data Management AID Washington, D.C 20523	Information Systems Div. (G-FIS/84) U.S. Coast Guard, 400 7th St.NW Wash., D.C. 20590	Division of Data Administration, NY State Dept. of Motor Vehicles Empire State Plaza Albany, N.Y. 12228	Office of Regulatory Info. Systems, Federal Power Commission 800 N. Capital Street, NE Wash., D. C. 20426	DoD Logistics Data Element Standardization & Management Office Room 7569 Hoffman Bldg.2 Alexandria, VA 22332	U.S. Navy Regional Data Automation Center (NARDAC) Bldg. 196 Washington Naval Shipyard Wash., D.C. 20374	Plans and Policy Div. Office of Automated Data Systems USDA Wash., D.C. 20250	Privacy & Data Adm. Division 810 Vermont Ave., N.W. Wash., D.C. 20420
CONTACT POINT	Data Base Administrator (202)254-8887	T. M. Kurihara (202)447-6261	Fernando Puente AUTOVON: 698-6001 Commercial (314)268-6001	Carolyn Moore (202)632-0084	Chief Information Systems Div. (G-FIS/84) (202-426-2438	Warren Crow (518)474-6871	John H. Yienger (202)275-4935	Aaron Hochman (703)325-9324	Frank Tagler (202)433-3571	Roxanne Williams (202)447-2617	Chief, Privacy and Data Adm. Division (202)389-3034
AVAILABILITY	On request	On request	On request	On request	On request	On request	Not available at present	Not available at present	On request	Not available	Documentation on request
TRANSFERRED	Yes	No	Yes	No	No	No	No	Yes	Yes	No	No

SYSTEM FEATURES - COMPUTER HARDWARE ENVIRONMENT

SYSTEM NAME	ADISDED	ASCSDBD	BARCON	DEMS	DLCS	DMVDED	FPCRIS	LOGORMS	STADES	USDADBD	VADD
MAINFRAME	UNIVAC SERIES 70/45	IBM 360/50	IBM 360/65	IBM 360/65	CDC 3300	MANUAL SYSTEM	IBM 370/158	IBM 360/65	HONEYWELL 6000	IBM 370/168	IBM 360/65
OPERATING SYSTEM	UNIVAC TODOS 23	IBM OS/MVT/HASP	IBM OS/MVT	IBM OS	MASTER		IBM MVS (VS2/EEL3)	IBM OS/MVT	HONEYWELL GCOS	IBM/VS2 TSO	IBM OS/MVT
CORE REQUIREMENT	60K	250K	150K	122K	20-45K (24 BIT WORDS)		50K	250K	29K	300K	90K
PERIPHERALS	1 DISK 3 TAPES 1 CARD READER 1 LINE PRINTER 1 PAGE PRINTER (OPTIONAL)	1 DISK 4 TAPES 1 LINE PRINTER	2 DISK 5 TAPES 1 CARD READER 1 LINE PRINTER	2 DISK 2 TAPES 1 CARD PUNCH 1 LINE PRINTER	1 DISK 4 TAPES 1 LINE PRINTER		1 DISK (WITH DYNAMIC ALLOCATION) 1 TAPE DRIVE 2 LINE PRINTERS	1 DISK 1 CARD READER 1 LINE PRINTER 4 DATA TERMINALS 3 DESK TOP PRINTERS	1 TAPE 3 TAPE DRIVES 1 CARD READER 1 LINE PRINTER 1 DATE TERMINAL	1 DISK 1 TAPE 1 LINE PRINTER OR TERMINAL	1 DISK 2-3 TAPES 1 LINE PRINTER 1 PAGE PRINTER (OPTIONAL)
OPERATIONAL MODE	BATCH	BATCH	BATCH	BATCH	BATCH UPDATE INTERACTIVE OR BATCH OUTPUT		INTERACTIVE OR BATCH	INTERACTIVE OR BATCH	INTERACTIVE OR BATCH	INTERACTIVE OR BATCH	BATCH
OTHER HARDWARE IMPLEMENTATION	HONEYWELL 66/80 IBM 360/370	NONE	IBM 360/370	NONE	NONE		NONE	NONE	UNIVAC 1100 IBM 360/370 CDC 6700	NONE	IBM 360/40

SYSTEM FEATURES - COMPUTER SOFTWARE ENVIRONMENT

SYSTEM NAME	ADISDED	ASCSDBD	DARCOM	DEMS	DLCS	DMVDED	FPCRIS	LOGDBMS	STADES	USDADBD	VADD
PROGRAMMING LANGUAGE	ANS COBOL	ANS COBOL	ANS COBOL	ANS COBOL	ANS COBOL	NOT APPLICABLE	MRI-SYSTEM 2000 ANSI COBOL	ALC/FORTRAN	ANS COBOL	MRI-SYSTEM 2000	ANS COBOL
USE OTHER PGMS	TDOS 23 UTILITIES	OS UTILITIES HASP	OS UTILITIES	UTILITY SORT	NO	NOT APPLICABLE	MEDIA TRANSLATE EDIT, LOAD AUDIT, OUTPUT GENERATION	NO	NO	TSO-V52 UTILITIES	UTILITY SORT UTILITY PRIN
USE DBMS	NO	NO	NO	NO	YES, CDC MARS-III	NOT APPLICABLE	YES, SYSTEM 2000	YES, CCA MODEL 204	NO	YES, SYSTEM 2000	NO
FILE ACCESS	SEQUENTIAL	OS SAM/ISAM	ISAM	SEQUENTIAL	RANDOM	WORD REFERENCE LIST	DIRECT	INVERTED LIST	SEQUENTIAL OR RANDOM	DIRECT/LIST	SEQUENTIAL
FILE STRUCTURE	SEQUENTIAL	SEQUENTIAL INDEX SEQUENTIAL	INDEX SEQUENTIAL	SEQUENTIAL	INVERTED LIST	ALPHABETIC	INVERTED LIST	DIRECT ACCESS LINKED LIST	INDEXED SEQUENTIAL OR DIRECT ACCESS LINKED LIST	LINKED/LIST	SEQUENTIAL
FILE BACKUP/ RECOVERY	YES	YES	YES	YES	YES	INDIVIDUAL DATA ELEMENTS HAVE BACKUP FOLDERS	SYSTEM 2000 FACILITY	YES, CCA MODEL 204	YES	SYSTEM 2000	BACKUP OF FILES IS A PART OF THE JCB STREAM

16

SYSTEM FEATURES - USER ENVIRONMENT

SYSTEM NAME	ADISDED	ASCSDED	DARCOM	DEMS	DLCS	DMDED	FPCRIS	LOGDMS	STADES	USDADBD	VADD
INPUT LANGUAGE	Fixed Form	Fixed Form	Fixed Form	Fixed Form	Fixed Form	MANUAL SYSTEM	Fixed or Free Form	Fixed or Free Form	Fixed Form	Fixed or Free Form	Fixed Form Transactions Free Form Description
COMMAND LANGUAGE	Limited DED/D Capability; Also Operating System JCL	Operating System JCL	Operating System JCL	Limited DED/D Capability; Also Operating System JCL	Command Language Written in Mars III		System 2000 Command Language	Model 204 USER Language	RAS III USER Language	System 2000 Natural Language	Operating System JCL
QUERY CAPABILITY	Batch, Predefined	Batch Predefined	Batch Predefined	Batch Predefined	Batch or Interactive Predefined and Ad Hoc		Batch and Online, Predefined and Ad-Hoc	Online Predefined and Ad-hoc	Batch and Online, Ad-Hoc and Predefined	Batch, and Outlines; Ad-Hoc and Predefined	Batch via Control Cards; Predefined
UPDATING TECHNIQUE	Individual Update Changes Only Specified Occurrence	Individual Update Changes Only Specified Occurrence	Allows Both Individual & and Single Entry Update	Individual Update Changes Only Specified Occurrence	Individual Update Flags Affected Entries		Individual and Single Entry Update	Online Create (add) and Change Simple Entry and Individual Update	Single Entry or Batch Individual Update	Individual Update Changes only Specified Occurrence	Individual Update Changes Only Specified Occurrence
OUTPUT FORM	Hard Copy	Hard Copy	Hard Copy Microform	Hard Copy	Hard Copy	Hard Copy	Soft Copy(CRT) Hard Copy	Soft Copy(CRT) Hard Copy (Printout) Magnetic Tape Microform From Magnetic Tape Using Com. Equip.	Soft Copy(CRT) Hard Copy (Printout) Magnetic Tape	Hard Copy	Hard Copy

DATA ELEMENT ATTRIBUTES DESCRIBED BY SYSTEM

SYSTEM NAME	ADISDED	ASCSDBD	BARCOM	DEMS	DLCS	DMYDED	FPCRIS'	LOGDMS	STADES	USDADBD	VADD
NAME	1-30 A/N CHARACTERS LEFT JUSTIFIED	1-60 A/N CHARACTERS LEFT JUSTIFIED	1-67 A/N CHARACTERS LEFT JUSTIFIED	1-40 A/N CHARACTERS LEFT JUSTIFIED	1-30 CHARACTERS	NO RESTRICTIONS	250 CHARACTERS LEFT JUSTIFIED	250 A/N CHARACTERS LEFT JUSTIFIED	700 A/N CHARACTERS LEFT JUSTIFIED 10 OCCURRENCES 70 CHARACTERS EACH	NO RESTRICTIONS	14 A/N CHARACTERS LEFT JUSTIFIED
Cobol Name	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No
Synonyms	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes
Character Type	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Length	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Justification	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No
Validation	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	Yes
Definition/ Description	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes
Relationship	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
User/Owner	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes

ELEMENT OCCURRENCES LISTED

SYSTEM NAME	ADIS	ASCSDDD	DARCOM	DEMS	DLCS	DMVDED	FPC RIS	LOGDMS	STADES	USDA DBD	VADD
Publication	No	No	Yes	No	No	No	Yes	Yes	Yes	No	No
Forms	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No
Program	No	No	Yes	Yes	Yes	No	Yes	Yes	No	No	Yes
Report: Manual	Yes	No	Yes	Yes	No	No	Yes	Yes	Yes	No	No
Report: Automated	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes
System	No	No	Yes	Yes	No	Yes	No	Yes	Yes	No	Yes
File (Existing)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes

19

OUTPUT/PRODUCTS

SYSTEM NAME	ADIS DED	ASCS DDD	DARCOM	DEXIS	DLCS	IDW DED	FPC RIS	LOGDBMS	STADES	USAMA DDD	VADD
Data Dict./Dir.	Yes	Yes	Yes	Yes	Directory	Yes	Yes	Yes	Yes	Yes	Yes
Reports: Summary/Stat.	No	Yes	Yes	Yes	---	No	Yes	Yes	Yes	Yes	Yes
Usage	No	Yes	Yes	No	Yes	No	No	Yes	Yes	No	No
User-Struct.	No	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	No
Change-Effect	No	Yes (Storage Structure)	Yes	Yes	Yes	No	Yes	Yes	Yes (By Display of Relationships)	No	No
Audit Trail	No	No	Yes	No	No	No	Yes	Yes	Yes	No	Yes
Redundancy Check	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	No
Error	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
MAIC/MADC	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	No
Indices	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No	No
Cross Reference	No	Yes (1)	Yes (9)	No	No	Yes (3)	Yes (5)	Yes (26)	Yes	No	Yes (1)
PDM. I/O Area	No	Yes	Yes ANSI COBOL	No	No	Not Applicable	No	Yes	No	No	No
PDM. Document	No	No	Yes	Yes	No	Not Applicable	Yes	Yes	Yes	No	No

APPENDIX B

SCOPE-NOTES FOR THE FEATURE LIST OF DATA ELEMENT DICTIONARY/DICTIONARY (DED/D) SYSTEMS

(The Feature List is designed for "yes/no" and short answers -- further explanation are found in the "Narrative System Descriptions" in Apprendix C).

GENERAL INFORMATION

SYSTEM NAME: Name or acronym by which the DED/D is known.

DEVELOPER: Name of organization that developed the DED/D.

ADDRESS: Address of developer.

CONTACT POINT: Name/Title and phone number of person(s) or office to contact for further information.

AVAILABILITY: Availability of the DED/D system to other government organizations.

TRANSFERRED: Has a copy of the DED/D been implemented in other government agencies?

SYSTEM FEATURES-COMPUTER HARDWARE ENVIRONMENT

MAINFRAME: The type of computer that the DED/D is implemented on, (e.g., IBM 360, CDC 6600).

OPERATING SYSTEM: The operating system under which the DED/D operates, (e.g., OS/360, Univac 1108 Exec II).

CORE REQUIREMENT: The amount of memory required to drive the DED/D.

PERIPHERALS: The peripheral equipment required to process the DED/D; (e.g. number of disks required, the number of tape drives).

OPERATIONAL MODE: Mode in which the DED/D operates, i.e., batch or interactive.

OTHER HARDWARE IMPLEMENTATION: Has the DED/D system been implemented on other hardware?

SYSTEM FEATURES-COMPUTER SOFTWARE ENVIRONMENT

PROGRAMMING LANGUAGE: Major programming language in which the DED/D

software is written.

USE OF OTHER PROGRAMS: Note required programs, which may be provided by the host system, needed to run the DED/D (e.g. sort-merge utility routine, a tape utility routine, etc).

USE DBMS: Does the DED/D require the use of a DBMS for its principal software?

FILE ACCESS METHOD: What file access method does the DED/D use? (e.g., Index Sequential Access Method, Random Access Method.)

FILE STRUCTURE: What file structure does the DED/D support? (e.g., Sequential, Direct Access, Linked List.)

FILE BACKUP/RECOVERY: Does the DED/D have its own backup and recovery capability? If not, indicate other means.

SYSTEM FEATURES-USER ENVIRONMENT

INPUT LANGUAGE: Does the input language require fixed-form or free-form transactions? Note other methods of input.

COMMAND LANGUAGE: Is the language a distinct capability of the DED/D, or does it rely on other means?

QUERY CAPABILITY: Is the DED/D query capability online or batch, predefined or ad-hoc?

UPDATING TECHNIQUE: What kind of file updating capability does the DED/D have? Is it "Single-entry" update, (the DED/D automatically updates every occurrence of the affected entry), or "individual" update, (only the specified occurrence is updated, and other occurrences have to be updated individually)?

OUTPUT FORM: Note the physical form of the DED/D system output, i.e., hardcopy (any printed output), soft copy (CRT).

DATA ELEMENT ATTRIBUTES

NAME: Provide restrictions placed on the name of the data element (e.g. length, character type, justification).

COBOL NAME: Does the system provide the capability to list the COBOL names of the elements documented?

SYNONYMS: Does the DED/D provide a word, an abbreviation or a phrase that may be used as a substitute for the data element name, conveying the

same meaning to the user?

CHARACTER TYPE: Does the DED/D have the capability of documenting the types of allowable characters (e.g., alpha numeric, special) used in expressing the content of a data element?

LENGTH: Does the DED/D have the capability of documenting the length of a data element?

JUSTIFICATION: Does the DED/D have the capability of documenting the justification of the data element?

VALIDATION: Does the DED/D document the validation (or edit) criteria, including value range and other non-quantifiable validation techniques for the data element?

DEFINITION/DESCRIPTION: Does the DED/D provide the capability of defining or describing a data element?

RELATIONSHIP: Does the DED/D provide the capability for specifying the relationship of a data element to another data element or to a higher level entity?

USER/OWNER: Does the DED/D provide the facility for indicating authorized users (persons or files) or owners of data elements?

DOCUMENTATION OF DATA ELEMENT OCCURRENCES

PUBLICATION: Does the DED/D document the publication(s) which prescribes the use of the data element?

FORMS: Does the DED/D document the forms which use the data element?

PROGRAMS: Does the DED/D document the computer programs which use the data element?

REPORTS: Does the DED/D document the manual and/or automatic reports that use the data element?

SYSTEM: Does the DED/D document the applications and/or systems that use the data element?

FILE: Does the DED/D document the files in which the data element appears?

OUTPUT/PRODUCTS

DATA DICTIONARY/DIRECTORY: Is the listing of the Data Dictionary/Directory one of the DED/D's formal products?

REPORTS:

SUMMARY/STATISTICAL: Does the system produce summary and statistical reports?

USAGE: Does the DED/D produce reports on user activity?

USER-STRUCTURED: Can users structure their own reports on an ad-hoc basis?

CHANGE-EFFECT: Does the DED/D produce reports that tell the effects a change in a data element(s) would have on all operational systems?

AUDIT TRAIL: Does the DED/D produce an audit-trail report (e.g., Who changed what element, when was the element created, when was the element updated or edited)?

REDUNDANCY: Does the DED/D produce a redundancy (multiply defined data elements) report?

ERROR: Does it produce an error (including inconsistency in data elements) report?

KWIC/KWOC INDICES: Does the DED/D have keyword-in-context (KWIC) or keyword-out-of-context (KWOC) indices capability?

INDICES: Can it produce other indices in addition to KWIC or KWOC indices?

CROSS-REFERENCE: Does the DED/D produce cross-reference reports, if so, how many types of cross-referencing reports can it produce?

PROGRAM I/O AREA: Can the DED/D produce I/O area file descriptions for computer programs?

PROGRAM OR SYSTEM DOCUMENTATION: Can the DED/D produce computer program or information system documentation?

APPENDIX C

Narrative System Descriptions

Associate Directorship for Information
System Data Element Dictionary
U.S. Civil Service Commission

ADISDED

ASSOCIATE DIRECTORSHIP FOR INFORMATION
SYSTEM DATA ELEMENT DICTIONARY
U.S. Civil Service Commission

I. Identification

The U.S. Civil Service Commission's Bureau of Manpower Information Systems (BMIS) installed its Associate Directorship of Information Systems (ADIS) data element dictionary (DED) during the spring of 1975. It is the responsibility of the Data Base Administrator (DBA). The phone number of the DBA is (202) 254-9603 and the address for correspondence is U.S. Civil Service Commission, BMIS DBA, Room 6424, 1900 E St. N.W., Washington, D.C. 20415.

For information, please contact:

U.S. Civil Service Commission
BMIS Data Base Administration
Room 6410 F
1900 E Street N.W.
Washington, D.C. 20415

Telephone: 202/254-8887

Documentation in the form of a brief operations manual, JCL, and an ANSI COBOL source list is available to interested government offices.

II. Scope

BMIS's purpose in implementing the DED system was two-fold. It was intended to document all elements so that for the first time CSC would have a comprehensive picture of how many elements it had and where and how they were used. Also, since the purchase of a new and much larger computer facility was planned, the directory could be used to analyze existing files and reports so that systems redesigned for the new computer could make optimal use of data resources. The system is intended to be both a flexible means of documenting present data use and an analysis tool for transition to a data base data management environment. It is basically a dictionary system, because its basic function is definition. However, it does pinpoint data element use in reports and files entities and to this extent can be considered a limited data directory.

III. ADP Resources

The ADIS data element dictionary is currently installed on CSC's Univac Series 70 (formerly RCA SPECTRA) 45's running under the TDOS 23 operating system. ADIS requires 46K bytes of main memory, a single disk drive and two tape drives (an additional tape drive may be used in lieu of the disks), a card reader and a printer. No special peripheral hardware devices are required but CSC uses a Xerox 1200 off-line printing system to prepare output products.

The system's six computer programs are written in ANS COBOL. All but the first of these, a preprocessor for input transactions written by CSC, were obtained from Facilities Systems Office (FACSO) NCBC Port Hueneme, California. The five basic programs consist of edit, update, and three report generation programs. CSC's implementation of this system is in the public domain and CSC would be happy to provide interested parties with its software.

In addition to the above software, a card to tape utility, a parameter driven sort, and a tape to print program are required to execute the ADIS DED.

The proprietary report generation and file management product, Easytrieve, from Pansophic Systems Inc. has been used with ADIS files to generate ad hoc reports.

IV. Content and Products

Several data entities are described by the ADIS DED. These are representation sets (groups of elements sharing common representations), data elements, data code tables, files, and reports. These entities are related in the following way: elements and representation sets are interchangeable, other entities subordinate.

The ADIS DED master file is a magnetic tape file sequenced by record type within data element or representation set I.D. There is no single master record which contains all data about a single data element. Rather, 11 separate fixed length record types, each of which contains different types of information and each of which may have 99 occurrences, are grouped behind a single data element record. Record format A describes representation sets and data elements and contains an ID number, title, standardization indicator, mnemonic, size, characteristic (alphabetic, numeric, alpha-numeric) code, type (elements collected on individual employees, on positions, or on aggregates of position or employee), a format code (indicating whether an element contains literal data or is coded), and a standardization status code.

Record format B describes files and indicates file name, number, location of the elements within the file, and whether or not the subject element is a key field or not. Record C covers synonyms for the element name. Record D allows up to 98 representations of a code set to be defined while Record E allows up to 5,800 characters of definition of a representation set or element to be entered. Record F, G, and H indicate element usage, edit criteria, and source respectively and Record I contains reports symbol, number and title for all reports in which the element appears.

There are approximately 2,000 data elements entered in the DED described by about 10,000 records or an average of 5 physical records per element. Access to these records is sequential. The various ADIS DED outputs are created by sorting records in various sequences involving record type and data element ID and then listing them through a master index program.

The chief product produced by the ADIS DED is a dictionary which lists for each element: its I.D. number; name; size; class; synonyms, acronyms; the titles and I.D. numbers, and location of the element within all files in which the element is used; a narrative description of the element; and its users. A remarks area is also available. The dictionary contains one page per element; pages are produced on an as changed basis. These pages are printed in I.D. number sequence and bound in a loose-leaf fashion so that when an element's information changes, its entire page can be easily printed and replaced.

Several indexes are also produced in addition to the dictionary. The first of these lists, for each report, its constituent elements. The second lists the same information for all files, with the elements listed both in location order and alphabetic order. A third index simply lists all element names in alphabetical order and a fourth lists the names in order of I.D. number. A fifth is available in mnemonic sequence. A compressed dictionary is also produced which, unlike the main dictionary, concatenates several element entries on the single page. A list of elements in alphabetic order shows all files on which each element occurs. The only remaining standard product is a transaction edit listing.

The use of Easytrieve for ad hoc retrievals has been mentioned before. The only difficulty experienced to date with this approach has been the fragmented DED file structure which does not lend itself to easy file manipulation.

V. Operational Consideration

The source of the data which has been entered into ADIS thus far has been system documentation where it existed, supplemented by face to face contact with the ADP systems analyst responsible for the various systems covered. The information gathering effort was phased in two ways, i.e., it not only proceeded on the system by system basis but on a field by field basis. In other words, the first phase aimed at collecting identification data about elements in selected systems. It was not until this effort was complete that other application systems and supplemental element and representation set information was coded and entered.

The personnel used to code the data entered came from both ADP and ADP user organizations. All shared a data management orientation and all work was coordinated by Data Base Administration. Working from existing documentation and through interviews with data users, the DRD data base was accummulated over a period of 9 months.

Data integrity was addressed by having it collected by a team of professionals who knew with consistency what information they needed. Further, after the creation of the first reports and files a second go round was undertaken as a review so that corrections could be made. Those persons interviewed initially were allowed to criticize the products and to indicate where changes should be made.

The ADIS edit program edits for completeness of information, not for validity. An effort will be made to establish more complete validity and relational edits. Back-up to files is provided through retention of father and grandfather master files and associated transaction tapes.

All ADIS DED update, report generation, and query operations are performed in batch mode, usually overnight.

VI. Security

Access to ADIS is monitored and approved by the Data Base Administrator, although the only current users are the Data Base Administrator, the Data Standardization Section, and an ADP user, the System Development Division of the Bureau of Retirement and Insurance and Occupational Health (BRIOH). All updates and requests for standard reports must be submitted through DBA at least during the system's present state of development.

No security restraints are imposed on the products and they are generally available. Subsets of the main files have also been created for special purpose query and report activity for submission outside of the DBA's area. The purpose of the security that exists is data integrity, not restriction of knowledge about content.

VII. User Characteristics

The present users of the directory are DBA and the Data Standardization Section and BRIOH.

DBA has been concerned since its establishment with the differing forms of the same data as used throughout the Civil Service Commission. DBA's primary use of the DED has been to identify and document where data is used and where it is inconsistent. The DED has already been extremely useful in identifying differing forms of the same data and each user of it.

The Data Standardization Section of the Manpower Systems Management Division carries responsibility for developing government-wide standards for personnel data elements. The ADIS DED has been used by them as a tool for developing and documenting the definition of these data elements and as an automated repository for the information needed to publish the Federal Personnel Manual (FPM) Supplement 292-1 which deals with data element standards for personnel data.

BRIOH has been using the DED as a means of documenting the data elements it uses in its own systems. It is developing a comprehensive picture of the data it uses in maintaining the Civil Service Retirement System.

Each of these organizations, DBA, DSS, and BRIOH, although they have differing perspective on data resource management, have found means to use a common system. At CSC, it is almost certain that system designers and DP management, especially of systems design activities, will also become heavy users of the ADIS DED.

U.S. Department of Agriculture

Agriculture Stabilization and Conservation Service

Data Base Directory

ASCSDBD

USDA ASCS DBD

I. IDENTIFICATION

The USDA Agriculture Stabilization and Conservation Service (ASCS) Data Base Directory (DBD) was developed in 1973 to support the development of the ASCS Integrated System Project. The development of the ASCS DBD was done under the direction of the:

Data Systems Division
Agriculture Stabilization and Conservation Service
U. S. Department of Agriculture 20250

Tom Kurihara of the Technical Resources Staff of the Data Systems Division is the contact point for information concerning the ASCS DBD (202/447-6261).

The ASCS DBD is no longer operational due to a change in Agency plans. The ASCS Integrated System Project for which the ASCS DBD was developed to support was called off.

The ASCS DBD is being described here in order to present the design features that were contained in the system. The original documentation includes a functional description, system and program specifications, user manual, and COBOL programs. Some documentation is readily available, for example, input forms and sample reports. More detailed documentation requests will also be met if practical. (Tom Kurihara: 202/447-6261).

II. SCOPE

The objectives and operational uses for which the ASCS DBD was developed in support of the ASCS Integrated System Project were to provide:

- A capability to record data requirements,
- A capability to collect, standardize, and control data attributes (such as data element size, class, COBOL name, etc.),
- A capability to support system analysis (including a KWIC Index capability),
  - * To identify commonalities (for the purpose of combining),
  - * To identify redundancies (for the purpose of singularizing),

* To identify system/subsystem inter-relationships,

* To enable statistical analysis of workload in conjunction with the Program Directory (The ASCS DBD provided volume information whereas the Program Directory provided frequency information).

. A capability for recording and controlling data element codes and their representative values,

* A capability to compute data storage requirements (in three forms: raw characters, packed numeric, and with a null-value compression),

* A capability to assist in developing system design and program specifications, and

* A capability to assist in coding or generating the Data Description Language (DDL) for a DBMS.

The ASCS DBD was developed in-house and was operational in 1974 and 1975. The ASCS DBD was a dictionary level implementation. During the period in which it was operational, the ASCS DBD maintained 6000+ data elements, 1200 + data aggregates, and 400 + logical files.

III. ADP RESOURCES

The ASCS DBD was operational on an IBM 360/50 and ran in a batch environment under OS. The programs were written in ANSI COBOL and used the OS access methods, sequential and indexed sequential access methods, (SAM and ISAM).

IV. CONTENTS AND PRODUCTS

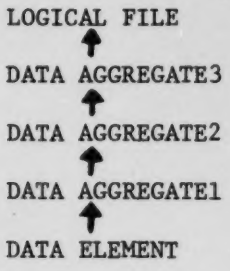
The data relationships described in the ASCS DBD were at three levels--data elements, data aggregates, and logical files. The three entity levels were fully integrated with ISAM chains connecting all entities. Definitions of the three levels were as follows:

. The data element was the lowest level of the ASCS DBD data base hierarchy and possessed a unique number, name, and other attributes. It was the smallest unit of a data string that could be accessed. (Example: ASCS FARMER IDENTIFICATION NUMBER).

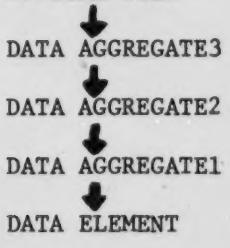
- The data aggregate was the middle level of the ASCS DBD data base hierarchy and possessed a unique number, name, and other attributes. It was composed of data elements, of subordinate data aggregates, or a combination of the two. (Example: ASCS FARM KEY--composed of 3 data elements: (1) STATE CODE, (2) COUNTY CODE, and (3) ASCS FARM NUMBER).

- The logical file was the highest level of the ASCS DBD data base hierarchy and possessed a unique number, name, and other attributes. It was composed of records containing data elements, data aggregates, or a combination of the two. A logical file could contain more than one record-type. (Example: ASCS FARM FILE).

The ASCS DBD was organized to express membership/ownership relationships in terms of residency and constituency. The ASCS DBD possessed chains moving upward to describe residency, i.e., the concept that a data element "resided" in a higher level such as a data aggregate or in a record of a logical file, e.g.,



The ASCS DBD also possessed chains moving downward to describe constituency, i.e., a logical file contained records which were comprised of "constituent" lower level entities such as data aggregates and data elements, e.g.,



The three master files used to house the ASCS DBD were stored in two versions--a sequential version used for updating and normal reporting and an indexed sequential version used for creating the return side of the chain, for checking for chain consistency, and for use in specialized reporting.

The ASCS DBD provided a variety of reporting support. The update segment provided a transaction register, edit and update error reports, and the three basic directories, the Data Element Directory, the Data Aggregate Directory, and the Logical File Directory. COBOL names for the three levels were generated by removing special characters, inserting hyphens, and truncating at 30 characters. A table of standard ASCS abbreviations was also used for substitution purposes. The logical maintenance segment provided the checking of chain consistency and the generation or removal of the return chain when new entities were either added or deleted. The analytical support segment provided a number of system analysis reports.

Extracts were made based on subsystems and segments within subsystem as well as at lower levels. Analytical reports were provided to denote inter-relationships among subsystems and segments. One report provided a ranking as to which logical files, data aggregates, and data elements were accessed the most. This report was provided by interfacing with the frequency information contained in the Program Directory. This report was particularly beneficial to the data base administration function, e.g., in the structuring of data and the assignment of keys. Another analytical report provided storage requirements of the data as specified in the ASCS DBD in several computational formats.

V. OPERATIONAL CONSIDERATIONS

The data collection effort was initially done by task forces, called Specification Teams (Spec-Teams), organized along subsystem lines. Each subsystem Spec-Team had a data specialist who was responsible for recording the data elements, data aggregates, and logical files needed to support the system requirements being specified for the subsystem. There were nine input forms used--a three-page set for each data element, a three-page set for each data aggregate, and a three-page set for each logical file. All data was entered via keypunching.

The overall coordination of the data specialists and the data specification effort was the responsibility of the data base administrator and his staff. This coordination function included the control and assignment of numbers for each data element, data aggregate, and logical file specified for the ASCS DBD. (The number included codes for breakout by subsystem and segment). Another

important function of the data base administrator was to identify redundancies and to combine or singularize as needed (for purposes of assigning the system and segment portions of a number, e.g., a data aggregate number, that was used by more than one system, the data aggregate would be assigned during the singularization process to a primary subsystem and segment).

The ASCS DBD interfaced with an ASCS Program Directory which contained information describing subsystems, segments, and computer programs.

The data in regard to computer programs included frequency information which when combined with volume information from the ASCS DBD was able to provide a statistical analysis of the workload.

VI. SECURITY

The ASCS DBD itself had no security provisions beyond the normal OS security capability provided for OS data sets.

The attributes that were collected for each data element and data aggregate did include a privacy/security classification as to which users had "read" authorization, which had "read/write" authorization, and which had neither authorization. The codes used to identify user organizations included a prefix to denote the "records-holding" user.

VII. USER CHARACTERISTICS

The user had an active and controlling role in the entire ASCS Intergrated System Project. It was during the User Requirements Specification phase that the data requirements were identified and specified for the ASCS DBD. These Spec-Teams were carefully chosen to include the various representative interests, and were chaired in all cases by user personnel.

The ASCS DBD was designed as a tool for ADP personnel and was not normally distributed to user Divisions.

VIII. COST

The ASCS DBD was developed in-house and maintained under the data base administrator function. It was operated in a batch mode in one of the Department of Agriculture's Computer Centers. The data collection for the ASCS DBD was initially done as a by-product of 14 Spec-teams that had been established during an early phase of the development cycle. Refinements, updates, and other changes to the ASCS DBD were accomplished as the system development continued. (The ASCS DBD was successful as a vehicle to establish change control of data attributes.)

Approximately eight system analysts and programmers participated part-time in the actual design, programming, and procedural development of the ASCS DBD system.

IX. GENERAL COMMENTS

The ASCS DBD experience suggests several conclusions:

Unquestionably, a DBD provides a convenient vehicle for collecting, standardizing, and effecting change control of an organization's data and its attributes.

For systems under-development, a DBD should be established early in a development phase to assist the designers of the system.

A DBD is an awkward and cumbersome thing when implemented in government agencies or large sub-agencies such as ASCS. The number, complexity, and multiple functional aspects of the many ASCS programs, many of a disjunct nature, causes data base management to be an immense challenge.

The ASCS DBD generated mountains of paper. (Future ASCS plans had been made to go to COM or CRT display output.)

Since commercial software products are typically marketed at a fraction of development cost, the ASCS experience would further suggest that purchasing of one of the available commercial packages would be a preferable course of action. In-house DBDs frequently have a tendency to never reach completion status. In the ASCS case, new capabilities were constantly being added until the system grew beyond recognition of at least one of its original designers.

Department of Defense - U.S. Army

U.S. Army Materiel Development
and Readiness Command

Data Element Dictionary/System Control and Documentation

DARCOM DED/SYSCAD

I. IDENTIFICATION

The US Army Materiel Development and Readiness Command (DARCOM) Data Element Dictionary (DED) and System Control and Documentation (SYSCAD) systems were developed by the US Army Automated Logistics Management Systems Activity (ALMSA) to support the development, documentation, fielding and control of a large, integrated, standard wholesale logistics management system identified as the Commodity Command Standard System (CCSS). The proponent for these systems is:

Commander
US Army Automated Logistics Management
Systems Activity
ATTN: DRXAL-ID (DED-Fernando Puente)
PO Box 1578
St. Louis, MO 63188

Telephones are AUTOVON 698-6001 or Commercial 314-268-6001.

System documentation is internally developed and maintained. The input documentation for the DED is DARCOM-R 18-5, Vol 4, Methods and Standards--Data Elements and Codes Standardization Program.

II. SCOPE

Early in the planning and development stages of the CCSS design, the challenge and need for data element standardization was recognized, particularly as it related to the integrated storage and file organization principles that were to be used. A single authoritative document for use as a communication link between the system designers, computer systems analyst, programmers, and subject matter specialists was recognized as being required. Out of these basic requirements, the DARCOM Data Element Dictionary/Directory system began to evolve. The first version of the DED system was implemented in early 1967. Its primary purpose was to produce an authoritative document for use by all functional and computer oriented personnel who were designing and programming the CCSS. The dictionary consisted of data element/data field descriptions that included standard names, programming mnemonics, definitions, automated data processing characteristics of the data, and the regulatory reference and authority for the entry. Also included was "where used" data, as it related to identifiable cells and subcells within the system which equated to major tasks and sub-tasks which were to be automated and integrated within the logistics system. The data elements that were entered in the DED data base were primarily developed by a committee of functionally oriented subject matter specialists, with the assistance of computer systems analysts. Control of the project was exercised by one organizational division which was charged with coordinating and controlling the total CCSS systems design.

The development of the basic data entries for the dictionary involved a great deal of coordination and discussion between the functionally oriented subject matter specialists. Many of the basic data entries that were developed crossed functional area boundaries and in some cases were referred to by differing names or had slightly differing contextual definitions. Before a data element entry was accepted for input into the DED data base, all functional representatives on the committee were required to sign off on the entry as being correct and required by the system. All entries entered into the DED were assigned a functional proponent who was given the responsibility for maintenance after it was entered into the data base. The DED that was produced became the communication link between all the personnel involved in the system development. Because the eventual operational CCSS was to be installed at six different commands in the eastern half of the United States, it also became a very important document to the eventual users of the system as they correlated data in their diverse existing files, to that which would be required for conversion to the standard system master file formats that would be required by the operational CCSS.

In the first design of the data element dictionary system, all mnemonics assigned to data element entries were limited to a maximum of 15 character positions. Through experience, this was found to be unsatisfactory, particularly when attempting to create meaningful programming mnemonics when documenting long names and attempting to maintain a semblance of standard abbreviations within the mnemonic. In a subsequent redesign of the data element dictionary system, the allowable length of a mnemonic was increased to 30 character positions to match the full capability of the COBOL language. In the CCSS, the maximum allowable length was increased to 24 character positions. Six positions were reserved to identify the file, sector and segment as a prefix to the mnemonics which are used to identify data fields within hierarchically structured master files of the CCSS. Since the CCSS was to be programmed in COBOL, it was recognized that an important tool for managing the system would be the mnemonic programming tags that would be used within the application programs and would be used to describe the data fields within the CCSS master files. Because application programs written against system master files would be required to use standard file descriptions which were resident on the system libraries, a great degree of discipline was imposed on the application programmer as he manipulated data within the master files or passed data to other application queues for use in other system processes. Because the programming mnemonic was to be embedded within the application programs and the master file descriptions, and to a certain degree reflect what was being manipulated within the system, the mnemonic was made the major key of the data element dictionary system, which in turn dictated that all mnemonics developed for entry into the data element dictionary data base must be unique.

The physical output format of the printed DED has remained fairly constant since its first publication. It contains an introductory chapter that explains the content of the dictionary. Next is a Key Word in Context (KWIC) index that contains all words used in the names assigned to the entries in the basic dictionary, listed alphabetically by the occurrence of all words used in all the names. The KWIC index is used as the primary research tool when trying to determine if a data entry has already been defined in the data base. Next is an alphabetically sequenced mnemonic index cross-referenced to the in-the-clear names that have been developed for the basic entries. The last section is the basic dictionary, sequenced alphabetically by the names of the data element entries. For the purpose of actually designing systems and for use as a functional definition reference, the sequence for this section has, in our experience, been found to be the most practical when seeking data definitions for either existing systems or for systems that are being designed. The format also allows for cross-referencing related entries within the dictionary to show hierarchal relationships. The ability to maintain these relationships is designed into the system.

Early in the dictionary system development, a decision was made to allow other system design activities to register as users of published dictionary entries or, if necessary, develop and become proponent for entries that had not been previously entered. This requirement led to the development of the capability to accept input from multiple remote activities and to register their interest in entries to the dictionary, or to accept submittal of new entries. Also developed was the capability to protect proponents of data entries in the dictionary from unauthorized update of this data. This was accomplished in the edit processes by developing system tables that cross-checked submitter input codes against entry proponent identifications in the data base.

To control and document the CCSS master files, a companion data base to the DED system was developed to act as a directory for the data that was being stored in the various master files of the CCSS. This portion of the system when fully developed was named the Systems Control and Documentation (SYSCAD) system. By coupling the two data bases, data element entries which were to be in the CCSS master files could be edited against the DED data base to insure that mnemonics that were being used in the CCSS master files descriptions had been established and that the automated data processing field characteristics were compatible with the DED entries. Since the SYSCAD master file contained all the CCSS master file descriptions, standard COBOL copy descriptions in both COBOL F and ANS COBOL were generated and established on the systems libraries for mandatory use of the application programmers when accessing any CCSS master file.

Well in advance of the installation of the CCSS, the eventual users had to know the structure and content of the CCSS master files so that they could begin correlating data from their system files to that which would be required for the standard CCSS file structures and data content. The file guide publications produced by the SYSCAD system were the primary documents used for this purpose. Also provided the user from the SYSCAD system, for use as an automated file conversion aid, were magnetic tapes that contained element by element descriptions of the master files for use in automated conversion methodologies. These magnetic tapes also contained change indicators so that the user could be kept informed of any changes that might be occurring to the master files. Also contained in the SYSCAD system are computed data starting positions for all data fields identified in the CCSS master file and computer segment lengths for both fixed and variable length records. The descriptions also identify all data fields that are used as file keys. Operationally, length and key position information is also used for keeping internal CCSS software control tables in synchronization.

The CCSS in its present configuration contains 29 master files. For the most part, files that were developed were designed to reflect the data required by the various functional logistics areas, taking into consideration the conservation of processing time, effective use of storage devices, ease of data maintenance and protection of data integrity within the system. These files in certain run configurations may stand alone or be coupled with other master files to run in a data sharing configuration. The CCSS has one primary file that contains common data that may be used selectively or in combination by any functional process in the CCSS. All other files may be influenced directly or indirectly by this file. The structure of these files is controlled and documented through the use of the DED system and the SYSCAD system. Since the first phase of the CCSS became operational in April 1971, the CCSS has continued to be changed and enhanced. This, of course, has included changes to the master files structure and content under a concept of scheduled system release management. In support of the release management concept, the SYSCAD system is designed to maintain, control and contain previous master file structures for regression test purposes, current file structures and future file structures as they will be in projected releases of the system. With the capability to manage future file structure releases, standard file description can be made available to the application programmer with sufficient lead time so that system changes may be coded and tested for whatever system release that the changes are required. This capability is very important because the magnitude, complexity and scope of a future systems change may require that work begin on a systems change many months before it is actually fielded.

The DED and SYSCAD systems when coupled, constitute a Data Dictionary/Resource Directory which has the capability to record data field descriptions, record their use within multiple media and in the case of files, their location and interrelationships. It should be noted that the DED system may be operated in a stand-alone mode as a dictionary with a limited directory capability. To exercise the control and documentation features of the SYSCAD system, it must be coupled to the DED system, it may not stand alone.

III. ADP RESOURCES

A. Hardware Requirements

1. IBM Model 360-65 or better.
2. 2314 disk unit or equivalent. Two required.
3. 2401-1 or equivalent 7-track tape drive, 800 BPI, if microfilm is to be produced for Stromberg-Carlson microfilm (Optional).
4. 2400-3 or equivalent 9-track tape drive, 1600 BPI. Five required.
5. 2501 card reader or equivalent.
6. 1403 printer or equivalent.
7. Core memory should be able to accommodate program sizes up to 150K, and an IBM Operating System (OS) version 21.8.

B. Software Requirements

1. All programs are written for ANS COBOL Version 3.1 Compiler.
2. IBM utility programs that are required are:
   a. IEBDG.
   b. IEBISAM.
   c. IEBGENR.
   d. IEHPROG.

IV. CONTENTS AND PRODUCTS

Every entry in the DED data base contains a unique name that can be a maximum length of 67 character positions. Each entry also contains a unique mnemonic which may be a maximum of 24 characters positions and must be constructed as a valid Common Business Oriented Language (COBOL) programming mnemonic. A 30 character position mnemonic may be entered with an override code. The documentation and programming standards for the CCSS require that this mnemonic be used both when documenting the system or within system application programs which use the COBOL language. Other required attributes of the data element entry are the characteristics of length, type and data field justification, the regulatory authority or reference for the entry, data field data security, the date the entry was last changed, the identity of the entry proponent, the definition, in which cells and sub-cells the element is used and if the element is locally derived, all of its data codes and items. Also provided is the capability to register other known references to uses of the entry related to local publication references, higher level regulatory references, report usages related to Report Identification Number (RINs), Report Control Symbols (RCSs) and forms usage. Application programs, file usage, systems usage references and known synonym mnemonic/abbreviations may also be registered. All of these references are available on automated cross-reference indexes for use in "as-required" inquiry processes.

The SYSCAD file contains information that describes the structure and data content of controlled system data base master files. Each data field in the master file being described must use mnemonic abbreviations established on the DED data base, against which the characteristics of length and data type are edited. The file descriptor entries also contain such information as the file acronym, record type, COBOL level, decimal positions, justification, key indicators, compiler to be used, occurrence or redefinition information, release number and information to be used for producing a file guide publication. The SYSCAD data base will only update after data editing has been successful against the DED data base. The SYSCAD data base when updated produces system master file COBOL Copy (both F and ANS) which is then available for use by the application programmers by whatever release version that is in work. File descriptions may also be prepared, which document the data content and structure of the file for documentation purposes. The DED file structure attached to this report is an example of a file guide description generated from the DED/SYSCAD systems.

The DED system contains 62 application modules. The system consists of the following tables and files on disk:

1 System editing table (May contain a maximum of 99 tables within it)

1 Master file containing--
  1 data base control record
  6 data record types

7 Cross reference indexes (inverted lists)

The files are Index Sequential, BISAM organization which can be read or updated sequentially or randomly.

The SYSCAD system contains 41 application modules. The system consists of the following tables and files:

1 System Master file on tape.

3 Master System Tables on disk.

A combination of tape sequential and random disk processing is used in this system. The tape file is processed sequentially with selected records being read to and updated on disk. Tables are Index sequential, BISAM organization which can be read or updated sequentially or randomly.

The DED System produces 34 reports which include formatted and selected listings of the data base and operational audit trail type reports.

The SYSCAD System produces 23 reports which include formatted and selected listings of the data base and operational audit trail type reports.

The reports are formatted using IBM Report Writer.

V. OPERATIONAL CONSIDERATIONS

Data for the update of the DED and SYSCAD data bases is controlled by one central activity. The DED input is received from any organization within the Command that is designing, programming or documenting automated systems. The input may be created by either functional or ADP specialists. The SYSCAD data base is updated by computer specialists whose function it is to control and document the system master files and their standard file descriptions.

All input is edited by system edit modules which do data field checks plus relational data checks prior to acceptance for data base update. The DED data base is designed so that only proponents of an entry may update the basic attributes of an entry, but other activities may register as users. After initial establishment of an entry, fixed field data may be updated independently. Both systems provide for periodic copying to tape in unload format for possible backup the recovery of the data bases.

Both systems operate in batch mode for updating purposes. The DED data base is updated monthly. The SYSCAD is updated on "as-required" basis. Inquiries into the DED are also in a batch mode on an "as-required"basis.

Both systems are used to control and document a large integrated wholesale Army logistics system known as the Commodity Command Standard System (CCSS) which includes processes in the functional areas of Financial Management, Provisioning, Cataloging, Stock Control International Logistics, Equipment Maintenance, Supply Management, and Procurement and Production. Other Command systems data fields are also registered in the DED.

VI. SECURITY

Both systems are designed to document unclassified data descriptions. Provision is made in the DED to record the security of the data that is being described. Also provided in the DED system is the capability to protect the basic integrity of a proponent's data field entry through the use of proponent tables.

Since these systems do not deal with "people" related data, the privacy act requirements have not been designed into the systems.

Data Elements Management System
Department of State - Agency for International Development

DEMS

62

AID'S DATA ELEMENTS MANAGEMENT SYSTEM

I. Identification

AID's Data Elements Management System was implemented in AID in July, 1974. This system, commonly called DEMS, is designed to produce a dictionary/directory of the data elements contained in the agency's information systems. The dictionary lists pertinent information about each data element, such as, data element title, definition, COBOL name, types and sizes, the programs, files and reports containing the data elements and, in some instances, the manual forms and reports where they appear.

The pilot phase of this undertaking was carried out between July and September 1974. Based on experience gained during the pilot phase, the system and procedures were evaluated and refined. Necessary revisions were made to permit more efficient coding, input, and analysis of data elements. As of December 31, 1975, approximately 93% of all existing ADP systems have been entered into DEMS. This represents in excess of 5,000 unique data elements.

Responsibility for the data elements management program, the Data Elements Management System (DEMS), and the output dictionary is vested in the Office of Data Management, Information Management Division. Mr. Linwood Rhodes is the Chief of the Information Management Division and Mrs. Carolyn D. Moore has responsibility for the Agency program, ADP System and Dictionary. She is also the contract monitor for a team of contractors responsible for support of the ADP system.

For information contact:

Mrs. Carolyn D. Moore
Information Management Division
Office of Data Management
Agency for International Development
Washington, D. C. 20523

Telephone: 202/632-0084

II. Scope

Scope and Objective - The objective of the Data Elements Management System was to produce a data elements dictionary/directory that provides a foundation for effective management of data contained in information systems. The dictionary produces listings of data tailored to the needs of various clients throughout the agency and provides a centralized location for all data, thus providing a basis for the evaluation essential to data integrity.

III. ADP Resources

The data elements management system is installed on the IBM 360/65. Its requirements are 122K core storage. All data is recorded on disk and converted to tape. There is one permanent disk pack using temporary space and one master disk. There are two output tapes and two tape drives. All files are recorded on backup tapes, and the system uses eleven (11) programs of which six are edit programs. The only software used is an internal utility sort.

The system's eleven computer programs are written in ANS COBOL. The system operates entirely in batch mode and update is maintained through punched cards. The update schedule is irregular.

IV. Contents and Products

The Data Elements Management system's primary report is the dictionary/directory report, which contains all the information in the master file and occurrence files. Such data as d.e. title, definition, COBOL name, type and size, and all the occurrences (programs, files, reports, systems) of the d.e. are listed. Several other reports may be obtained:

1. A listing of only the data elements and definitions, types and sizes and COBOL names. (This is useful when analysts are developing systems and need a shopping list of what's available.)

2. A listing by program or file or report ID of all the data elements used in that entity. (This too is an analyst aid.)

3. A listing, by (manual) form number or report number of all the data elements appearing on that form or report. (This is a boon to reports/forms analyst.)

4. A listing, by system ID, of all the data elements in a given system. This is good for management interests.

5. A listing of a single data element identified with all the occurrences. (This is used to evaluate the effect of changes.) Also, the system produces other miscellaneous reports as required.

Two data files, the master file and the occurrence files are maintained on disk packs. The master file has 500 characters for each record - the occurrence file has 102 characters for each record. There are several thousand records in each file. The files are in numerical sequence by data element code, however, the data elements may be sorted in alpha sequence.

Presently, there are approximately 5,400 unique data elements in the system with an approximate total of 84,000 lines printed. The capacity of the system is unlimited.

V. Operational Considerations

Analysts assigned to the maintenance of a system are notified when recording of the system's data elements has been completed and a report listing the elements and related information is provided for the analyst's use. At that point, maintenance on the data elements officially begins and the Information Management Division is notified in writing (Form 320-26), whenever data elements in that system are revised, added or cancelled. The participation of DM analysts is vital to the maintenance of systems presently included in the data elements dictionary, since a data element revision, deletion or cancellation in an ADP system greatly affects DEMS. For example, if the "type" and "size" of a data element are changed, then the information provided by DEMS, and all the files, programs, reports and systems which contain that data element in that particular application require revision.

The information is collected on input forms after it has been researched and checked by the recorders. AID is presently using five data technicians to record and maintain data contained in the system.

The data undergoes several internal checks before being processed. AID has achieved a two percent error rate, which is justified by the sheer volume of the data. The errors that do occur usually involve a misspelled word or an ommitted comma and do not affect the correctness of the data contained about a data element.

The edit program checks for conformity of data but does not actually validate the data. Backup files are maintained and punched cards are held for a reasonable length of time.

However, it must be stated that a system is only as good as its source documentation. Occasionally, the system's books or manuals from which data is extracted are incomplete, antiquated, or difficult to piece together.

Retrieval of data is through batch mode using keypunch cards.

VI. Security

There is a great degree of informal control exercised by the system's monitor. All input/output requests are routed through the monitor. Although the actual collection (in machine runs) is secured (limited official use), there is no control of the data once distribution has been made.

VII. User Characteristics

The data elements dictionary provides a logical starting point for identification and evaluation of data for purposes of data base management, conformance to Privacy Act requirements and security of data.

DEMS provides SER/DM analysts with a credible source of information about data elements contained in ADP systems documentation as follows:

(a) provides DM analysts with an organized aggregation of data elements and related information.

(b) serves as a catalogue to assist DM analysts in identifying data already available when new systems are being developed.

(c) provides for identification of areas affected when data elements are changed.

(d) provides the basis for usage of data that is acceptable throughout the agency, regardless of the application, thus promoting consistency and clarity in usage.

(e) permits forms and reports analysts to more effectively identify and control the data in agency information systems.

(f) provides a basis for forms and reports analysts to spot areas of redundancy and indeterminate data.

DM's clients benefit from that portion of the data elements dictionary which lists data elements' titles, definitions, types and sizes and COBOL names, as this information promotes consistency and effective interchange of data among the users of ADP systems. The dictionary is written representation of the data processed by DM for agency use.

Because effective management of data is built entirely on the concept of standardization of data (and its use) strong emphasis is being placed in this area.

Approximately 75-100 data elements are standardized and absolutely prescribed for use. This effort was quite difficult due to the politics surrounding data. However, progress is being made. Several of our data elements are used interchangeably among systems, and these data elements provide the foundation for out standardization efforts.

VIII. Cost

Costing information in the system, in terms of present resources, can be obtained by calling the project monitor.

Cost feasibility in terms of long-range benefits or improvement in "management of data" has not been realistically determined.

Data Label Control System

Department of Transportation - U.S. Coast Guard

DLCS

DATA LABEL CONTROL SUBSYSTEM

I. Identification

In 1972 the U. S. Coast Guard developed the Data Label Control Subsystem (DLCS) to aid in the development and maintenance of the Joint Uniform Military Pay and Personnel Accounting System (JUMPS). System specifications, program documentation, user procedures, and additional information are available from:

Chief, Information Systems Divsion (G-FIS/84)
U. S. Coast Guard
400 7th St. S. W.
Washington, D. C. 20590
Telephone: 202-426-1686

II. Scope

The DLCS was designed to act as a directory for the expected 285 programs, 580 computer records, and approximately 2000 data elements needed for the JUMPS system. The system provided a means to verify the correct use of common representations in the programs to be writen as well as providing a map of programs to be changed if an element or its code structure changed. The DLCS also provided a chance to test the Data Base Management System (DBMS) to be used in the JUMPS system. The system contained standardized names, field sizes, and other attributes of each element. A manual definition book is maintained to further define those data elements found in the DLCS.

III. ADP Resources

The DLCS was developed using the CDC proprietary DBMS, MARS-III on the CDC 3300 computer. Disk storage on one class "B" pack is needed for the three data files. Tape files are used for input and backup while output is to a printer.

A number of application programs were written to edit, update, dump, and restore the files. These jobs require 40 to 90 quarter pages of core consisting of 512, 24-bit words. A number of reports are produced using the MARS-III inquiry language, QUERY. These jobs require from 40 to 60 quarter pages of core.

IV. Content and Products

Data is organized into three MARS files, LAB1, LAB2, and LAB3 located on a single class "B" magnetic disk pack. Characteristics of each data element are common to all three files and appear in the "fixed length portion" in MARS terminology. The files are sequenced on an arbitrarily assigned sequence number (SONBR) for each data element. Each file has one or more variable portions, each of which consists of a variable number of fixed length areas. LAB1 contains non-standard COBOL abbreviations for each data element and a list of programs using the data element. LAB2 contains lists of input forms, reports, files, and records in which the data element appears. Transaction codes and associated function codes are contained on LAB3 for each data element.

Standard reports include an update and exception report for each file as it is updated. Reports of data element characteristics and usage are produced through the use of the inquiry language. Special inquiries are run on an as required basis.

V. Operational Consideration

Data is to be prepared by the systems analyst during the analysis and design phases of the new system development. Coding sheets in the necessary record formats are pre-printed. After keypunching, jobs are submitted in batch mode to update the file and produce standard reports. Backup dumps are taken on a regular schedule.

The data is minimally edited before update and errors encountered are listed on an exception report. Corrections are made by resubmitting the input forms.

A special means of updating the file and record lists is provided by one of the programs. This program reads file descriptions and record descriptions from a COBOL program, validates the COBOL names and updates the corresponding lists for that element.

DLCS

VI. Security

Users of the Data Label Control Subsystem must identify themselves to the Multiple Access Retrieval System (MARS) which processes the DLCS. User identification and passwords are both required for access to MARS, which is operated on the CDC 3300 computer system. However, there is no individual data element protection built into the DLCS, although the capability exists. Users of the system must rely on proper user identification and passwords to protect access to data.

Data security is further controlled by access lists of personnel authorized to process programs on the computer system. If a person submits a job without the proper authorization to run that program, the job will be aborted.

VII. User Characteristics

The DLCS is used by technical personnel (computer systems analysts and computer programmers). In addition, non-technical (user) personnel are involved with defining data element characteristics. The DLCS coordinator reviews *and* submits updates on an as required basis. Inquiries are prepared by the person desiring the material and submitted directly to the Transportation Computer Center, TCC.

VIII. Costs

The DLCS was developed by five persons in 1971 and 1972. The skills employed were those of computer system analysts and computer programmers. The total development time was approximately 24 months with an estimated cost of $30,500.

Department of Motor Vehicles Data Element Dictionary

State of New York Department of Motor Vehicles

DMV/DED

72

STATE OF NEW YORK
DEPARTMENT OF MOTOR VEHICLES
Data Element Directory (DMV/DED)

I. Identification

The New York State Department of Motor Vehicles is developing a Data Element Directory (DMV/DED) of its computerized files. The Data Standards and Control Bureau of the Division of Data Administration is responsible for producing and maintaining the directory.

For information, contact:

New York State Department of Motor Vehicles
Division of Data Administration
Empire State Plaza
Albany, New York 12228

Attention: Warren Crow
Telephone: 518/474-6871

The scheduled completion data for the directory is July 1977.

II. Scope

The Division of Data Administration was organized in 1972 to coordinate the management and protection of the Department's data resources. The Data Standards and Control Bureau of the Data Administration Division was directed in 1974 to develop and maintain a Department of Motor Vehicles Data Element Directory. The directory is intended to be a central data reference guide. It will be used to standardize data in the Department of Motor Vehicles thereby promoting data interchange among organizations involved in traffic safety.

Since the bulk of the DMV/DED contains definition information it is basically classified as a dictionary although it does contain some characteristics of a directory. The data elements are linked to records, files, transactions and their input, file and output representations are documented.

III. ADP Resources

The DMV/DED is not automated.

IV. Contents and Products

The DMV/DED contains three categories of information: identification, definition and representation. Identification information furnishes the data element with unique labels and links the elements to records,

files and transactions. Definition information explains the content, purpose, source and use of the data element. Representation information describes the length, type and format of the data item. There are five sections to the DMV/DED:

a) Explanatory Text
b) Data Element Description
c) Files Inventory
d) Transactions Inventory
e) Word Reference

A. The Explanatory Text describes the purpose, content, scope and use of the directory.

B. Data Element Description. This is the core of the DMV/DED. Descriptions are listed alphabetically by data element signature. Each data element description contains the following.

1. Signature - the complete unique name of the data element.
2. DMV Code - a four position alpha numeric code assigned to the data element for use in manipulating data.
3. Name-in-context - description of how the signature appears on forms where uniqueness is provided by context. (e.g., the data element "name, licensee" appears as "name" on the license certificate.)
4. Abbreviation - the shortened form of the signature as used on reports and forms.
5. Technical definition - explanation of the content and purpose of the data element.
6. Input source - identification of the originator of the data element.
7. Use - explanation of the general uses of the data element.
8. Special EDP use - description of the data element's use *for* specific data processing purposes. (e.g., access key, sorting field).
9. Synonyms - words similar or identical in meaning to the signature.
10. Computer files - list of files where the data element appears.
11. Transactions - list of transactions where the element appears.
12. Systems - list of sets of computer processes where the data element is used.

13. Components of a composite data element - if a data element consists of two or more other data elements it is referred to as a composite data element. The data elements that make a composite are referred to as components. For example, "date of birth" is a composite data element whose components are "year", "month" and "day".
14. Traffic Records reference - signature of the data element as it appears in the New York State Traffic Records Data Dictionary.
15. NHTSA reference - name and reference codes of the data element as it appears in the National Highway Traffic Safety Administration's Design Manual for a State Traffic Records System.
16. Comments - information that does not fit into any other category.
17. Representations

    a. Level - input, output, or file representation.
    b. Type - name, abbreviation, code or numeric value.
    c. Length - number of positions or bytes.
    d. Format - alphabetic, numeric, etc.
    e. Data item description - describes and illustrates the possible values of the data element.

D. Transactions inventory. For each transaction the following information is given:

1. Transaction name.
2. Description - a brief narrative of the purpose of the transaction.
3. System - a list of systems where the transaction is used.
4. Data elements. A list of data elements that appear in the transaction.

E. Word reference. Words in the DMV/DED are defined as signatures, names-in-context, abbreviations, synonyms and components of composite data elements. All words are listed alphabetically in the word reference and are linked to data element signatures. This was created to help users locate a particular data element description without knowing the full signature.

V. Operational Considerations

Our main source of data was system design documentation. Where the documentation was missing or incomplete, interviews with the analysts and programmers and some users provided the necessary information.

Our data collection tool is a form designed by our bureau based upon material developed by the ANSI Data Directory Committee (D-20.1), an outgrowth of the ANSI State Model Motorist Data Base Committee (D-20). All data element descriptions are to be cleared by pertinent analysts, programmers and users.

VI. Security

Some confidential restricted information such as personnel file transactions has been omitted from the directory but there are no plans for security restraints of the directory itself.

VII. User Characteristics

The anticipated heavy users are systems analysts, programmers and Data Standards and Control Bureau analysts. We also anticipate usage by DP Management, some users and other organizations involved in traffic safety.

Federal Power Commission Regulatory Information System
Federal Power Commission

FPCRIS

FEDERAL POWER COMMISSION
REGULATORY INFORMATION SYSTEM
DATA ELEMENT DIRECTORY

I. IDENTIFICATION

The Federal Power Commission, Office of Regulatory Information Systems, Data Management Branch, directed the development of the Data Element Directory as an ordered collection of data elements data items, source, usage and output references, and other inter-relationships to support its Regulatory Information System (RIS).

For information contact:
FEDERAL POWER COMMISSION
Office of Regulatory Information System
825 N. Capitol St. N. E.
Washington, D. C. 20426

ATTENTION: John H. Yienger
PHONE : 202-275-4935

FPC's Regulatory Information System and its associated Data Element Directory are currently under development. Detailed documentation is thereby not available at this time.

II. SCOPE

The Data Element Directory (DED) is a software-user tool through which the RIS data bank is managed. In addition to information necessary to describe data bases, data elements and data items, the DED has control information necessary to completely direct the operation of the generalized software necessary to acquire, edit, audit, manipulate and administer all the data bases of the RIS data bank. The DED performs the functions of a directory and a dictionary. In essence the DED is a data resource repository for the RIS.

III. ADP RESOURCES

The DED is installed as a system 2000 (S2K) data base on the IBM 370/158 operating under MVS. Other than the standard overhead, the DED has over 6 million characters of data. The DED can be processed interactively using System 2000 natural language or in batch mode using ANSI COBOL with System 2000 Programming Language Interface (PLI).

IV. CONTENTS AND PRODUCTS

The DED is a System 2000 Data Base itself and is essentially a three limbed tree. One limb is for the Data Bases of the RIS, in which the Data Elements are defined and cross-referenced in input documents, data item lists, or units-of-measure. Also the data element's input usage and output requirements relative to edits, audits and reports are specified. The second limb of the DED is the data item lists for the appropriate data elements. Each data item list is described and standardized by name, code and abbreviation. All the data item lists form a data register. The third limb of the DED is a list of the appropriate source documents which are the collection media employed to collect data. The titles of the fields on these schedules are the data elements.

V. OPERATIONAL CONSIDERATIONS

The Regulatory Information System and its associated Data Element Directory have been developed concurrently by the Federal Power Commission and its contractor, Planning Research Corporation. All data elements, data item lists, data bases and Forms/Schedules (source documents) are a product of an integrated development effort.

There is a unique set of source documents used to supply all the data required by the Data Element Directory. These input documents are prepared by Data Base Administrators, Data Standards Analysts, and Forms Designers. These documents collect all the definitions, characteristics, and interrelationships data contained in the DED. This data is coordinated and processed by the Data Base Administrator for the Data Element Directory.

All initial loading and updating of the DED are performed in batch mode for audit and control purpose, usually overnight. However, the DED is maintained on-line and is available for interactive queries and print-outs to support special analyses. No data base changes, to either the DED or the other RIS data bases are performed interactively. Under the direction of the DED approximately 2 billion characters of RIS data are collected, edited, audited, approved and maintained on-line for any 3-year period. The data are massaged for analysis via the constraints of the DED. This includes the DED's role as being the repository to define the parameters and other control information employed by the software to maintain the RIS.

VI. SECURITY

Security for the RIS and its DED are imposed for the protection of the data within its environment against accidental, passive and willful destruction or abuse. Presently, the security features are those provided by the IBM MVS operating system and system 2000 data base management system. These include several modes of password protection of the data. However, FPC's security program is continually under review.

VII. USER CHARACTERISTICS

The DED is developed, maintained and operated by the Data Base Administration group. This group is responsible for all data bases of the RIS and is in constant contact with users. Through this contract they are familiar with User requirements, responsible for RIS data base and source document design, and load, update and maintain the RIS data bases.

The primary users of DED are the Data Base Administrators. They use the basic reference and cross-reference features of the DED when designing data bases and source documents and in evaluating requested modifications or enhancements. In additon, the DED is used to define the data elements and their prescribed data items for filling out source documents.

VIII. CONCLUSION

The DED of the FPC's RIS is a highly specialized data resource manager tailored specifically for FPC's generalized data base software processing system. The DED is a System 2000 data base itself and is processed and maintained by the same generalized software system as all other RIS data bases. These factors must be thoroughly investigated when considering the possible use of the FPC's DED.

Logistics Data Resource Management System

Department of Defense - Office of the Secretary of Defense

LOGDRMS

I. IDENTIFICATION

Office of Prime Responsibility

The Department of Defense Logistics Data Resource Management System (DoD LOGDRMS) was developed by the DoD Logistics Data Element Standardization and Management Office (DoD LOGDESMO) to support execution of the Department of Defense Logistics Data Element Standardization and Management Program (DoD LOGDESMAP). For information on this system contact:

Chief
DoD Logistics Data Element Standardization
and Management Office (LOGDESMO)
Attention: Mr. A. Hochman
Room 7S69, Hoffman Building 2
200 Stovall Street
Alexandria, Virginia 22332

Telephones are: AUTOVON - 221-9324 or
Commercial - (202) 325-9324

Documentation

System documentation is internally developed and maintained. A procedures manual for the DoD LOGDESMAP has been compiled and is currently being coordinated throughout the Department of Defense prior to final approval, publication, and implementation. Additionally, a users manual is being compiled as a companion document to the procedures manual. Both documents are projected for issuance by June 1977.

II. SCOPE

Reason for Implementation

* The Department of Defense (DoD) Logistics Data Resource Management System (LOGDRMS) was developed to provide a more effective vehicle for the meaningful standardization and management of data employed within the DoD logistics community.

* LOGDRMS specifically is designed to provide:

** Uniform identification, categorization, and classification of the data employed in DoD logistics data systems.

II. SCOPE (Cont'd)

** An organized means for family grouping of related data representations and subjecting such groupings to the formal processes of standardization, i.e., simplification and the development, coordination, and publication of standards.

** A vehicle for storing, displaying and/or publishing standards to personnel engaged in data system management, development, design, operation, maintenance, or use in order to maximize reutilization of existing data as opposed to time consuming and costly generation of new data.

** Necessary information support services to management and operating elements so as to maximize their visibility of existing data and information and thus effectively improve their decision making processes.

** An analytical tool for recording the hierarchical relationships of data to the entities of which they are a member (e.g., forms, reports, files, etc.) and to the entities into which they logically subdivide (e.g., data elements, data chains, data codes, etc.) which can be utilized to assess the impacts of planned or proposed data system changes or in conducting required analyses.

** Critically needed management of logistics data as a "resource" in a manner similar to that employed in the management of other resources such as materiel, personnel, facilities, money, etc.

Background

* Department of Defense (DoD) Directive 5000.27, "Logistics Data Element Standardization and Management Program (LOGDESMAP)", dated March 28, 1975 (a) establishes, defines the objectives of, and assigns responsibility for the DoD LOGDESMAP; (b) establishes and defines basic principles and policies for the management of logistics data within the DoD; and (c) authorizes publication of a LOGDESMAP Operating Manual prescribing uniform procedures to be applied by participating DoD Components.

* DoD Directive 5000.27 identifies the objectives of the DoD LOGDESMAP as follows:

II. SCOPE (Cont'd)

** Improve the management of DoD resources at every level through the reduction of unnecessary overlap and duplication in DoD logistics data systems.

** Reduce the costs of DoD data systems design, development, programming and operation.

** Assure the effective management of logistics data bases.

** Facilitate the interface and integration of logistics data systems and the interchange of data between and among such systems.

** Provide a common base of standard data elements and related features for use throughout the DoD logistics community as well as other affected DoD communities.

** Permit more efficient determination of the impact of anticipated, proposed, and/or approved changes by those organizational elements planning, administering, and operating programs and systems employing logistics data.

* The following policies are enunciated in DoD Directive 5000.27:

** DoD Components will jointly develop uniform methods, procedures, and controls for the management of DoD logistics data bases.

** Management of logistics data on a system oriented basis (i.e., how data are used as opposed what data are) will continue to be accomplished by the organizations responsible for logistics data system development and configuration management.

** DoD standard data elements and related features will be developed and used unless specifically exempted by the Assistant Secretary of Defense (Installations and Logistics) ASD(I&L) in (a) the development or redesign of automated logistics systems and (b) authoritative issuances which prescribe the collection, reporting, or interchange of logistics data.

II. SCOPE (Cont'd)

** Issuances of a joint nature involving requirements for the development, identification, or interchange of logistics data will be coordinated under the provisions of the DoD LOGDESMAP.

Applicability

* The DoD LOGDESMAP applies to all DoD Component organizations involved in (a) the development, design, or modification of automated logistics data systems; and (b) the initiation, coordination, publication, or revision of authoritative issuances which require the collection, reporting, or interchange of logistics data.

* For the purpose of this program, the area of logistics encompasses all responsibilities assigned to the Assistant Secretary of Defense (Installations and Logistics) ASD(I&L).

History

* The improvement of DoD logistics systems has necessarily been closely related to the emergence and advances in the technology of automated data processing and improved communications. As the technology has progressed, the development of automated logistics systems has moved with varying degrees of speed and has been hard pressed to keep pace. In part, this has been due to the inability to capitalize upon existing technology; more importantly it is the result of long lead time necessary for effecting system changes. As component systems become progressively complex, these lead times have become increasingly prolonged. The final product is a series of subsystems or processes, each uniquely complex and difficult to change.

* This experience highlighted a critical need for a master plan aimed at preventing the proliferation of non-standard, incompatible, non-comparable and independently developed systems. Such a plan, identified as the DoD Logistics Systems Blueprint was promulgated in 1968 and implemented in 1972 as the Department of Defense Logistics Systems Plan (LOGPLAN).

* A vital part of the Blueprint addressing the standardization and management of logistics data was implemented in 1970 through initiation of the program ultimately formalized by DoD Directive 5000.27 (reference para. IIB).

II. SCOPE (Cont'd)

* The Blueprint portion dealing with logistics data pointed out that the rapid proliferation of logistics data language vocabulary variation resulting from independent system development actions in today's highly interdependent logistics environment seriously hampers effective accomplishment of logistics functional missions, particularly from an overall DoD point of view. As a consequence, each operating data system must maintain a built-in capability for translating the data language vocabulary of other data systems with which it interfaces. In order to overcome this continuing trend, it is intended that management discipline will be applied to data and information in a manner similar to that employed in the management of materiel. Data and information, like materiel, are susceptible to the full range of logistics life cycle functions (e.g., requirements determination; identification; standardization; acquisition; storage; distribution; utilization; disposal).

* Early efforts by LOGDESMO were manual, followed by a computerized system in batch processing mode. Considering the enormous volume of records and the extensive time required to update and publish information, a major effort to develop an economical effective improved system was undertaken.

* As a part of this effort, a test of the Navy's Record Association System II - Interactive (RAS III) was conducted. While it was concluded that RAS III was a considerable improvement over the constrained batch processing, a number of required improvements were evident. As a consequence, a comprehensive system functional requirement was developed. Based upon this requirement, a decision was reached to employ Computer Corporation of America Model 204 (DBMS software package) on a lease basis from the Department of Commerce and to use the leased software on that Agency's computer on a time sharing basis.

* More recently, an acquisition action to secure a software package (DBMS) capable of meeting LOGDESMAP requirements as well as other current or projected user requirements was undertaken by the Defense Logistics Agency (DLA). When such acquisition is completed, it is planned to terminate the lease agreement with the Department of Commerce and to employ DLA's computer located at Cameron Station, Alexandria, Virginia

Implementation Level

* LOGDRMS is a Department of Defense-wide system insofar as content is concerned.

* The system currently operates at three physical sites.

** DoD LOGDESMO - Updates (both create and change) are entered using data terminals. Query is conducted on a planned or ad hoc basis with print out optional on desk top line printers.

** Defense Administrative Support Center (DASC) - DASC (a field organization of the Defense Logistics Agency) provides programming of "segments" and testing and debugging using a test base. When approved, segments are entered in LOGDRMS data base for use by LOGDESMO personnel. All efforts of this nature are performed on-line using data terminals. DASC also provides batch processing services using magnetic tapes dumped from the physical data base. DASC also arranges for micromation support services to develop microfiche of LOGDESMAP publications derived from batch processed magnetic tape entered into Computer Output Microfilm (COM) equipment.

** Department of Commerce, Washington, DC is the owner of the Model 204 software and the computer employed for LOGDRMS. The physical files of LOGDRMS are stored at Commerce on a disk pack.

* Follow-on planning provides for the termination of the Department of Commerce lease arrangement with the DASC assuming all responsibilities formerly performed by Commerce.

* Additional long range planning provides for multiple update and query sites to be located at various system design centers throughout the United States.

III. ADP RESOURCES

Hardware Requirements

* IBM Model 360-65 (Department of Commerce)

* One (1) disk pack

* One (1) line printer

* Five (5) data terminals

** One at DASC

III. ADP RESOURCES

** Four at LOGDESMO

* Four (4) desk top line printers

** One at DASC

** Three at LOGDESMO

* One (1) card printer

* Batch Processing Hardware IBM 370-155 involving 20 magnetic tapes.

Software Requirements

* Computer Corporation of America Model 204 Data Base Management System (Department of Commerce)

* Program segments written in Model 204 User language

* Batch processing programs are written in ANSI COBOL

IV. CONTENTS AND PRODUCTS

Data Organization

* Data Bases: The LOGDRMS data bank consists of 25 data bases including one for:

a. Each Department of Defense Component

b. Other Federal Agencies (U.S. Government)

c. Interagency Groups/Committees

d. Non-government Organizations

e. State/Local Government Organizations

f. International Government Organizations

g. DoD LOGDESMO

* Sectors: Each data base subdivides into sectors as follows:

a. Organizations

b. Functions

IV. CONTENTS AND PRODUCTS (Cont'd)

c. Subject Matter

d. Issuances/Publications

e. Management Plans

f. Management Programs

g. Management Studies

h. Management Systems

i. Management Subsystems

j. Management Operations

k. Management Procedures

l. Automated Data Processing Systems

m. Automated Data Systems (Applications)

n. Data Systems (Processes)

o. Automated Programs

p. Data Files/Bases

q. Records/Segments

r. Formats

s. Forms

t. Documents (Input)

u. Reports (Output)

v. Data Fields/Blocks

w. Data Element Applications

x. Data Chains

y. Other Multiple Data Element Representatives

z. Data Element Categories

IV. CONTENTS AND PRODUCTS (Cont'd)

** Data Elements

** Data Items

** Terms

** Abbreviations

** Correspondence (LOGDESMO internal)

** Library Control (LOGDESMO internal)

* Logical Records - An array of physical records to logically represent a record within a sector of a data base. Each such logical record includes attribute information concerning the subject of the record. These attributes are logically arrayed in seven sections as follows:

** Record Identification Attributes including identification of the data base, the sector, the proponent organization, reference document, document subdivision, and date of latest record change together with a unique six position numeric Record Identification Code (RIC).

** Identification Attributes such as Official Name, Mnemonic Abbreviation, or Initialism, Document reference designation, synonymous names, standardization status information, subject matter identification, system configuration control designation, scope, version data, etc.

** Representation Attributes such as type of representation, recording mode, length in characters, type of characters, COBOL picture, signed quantity indicator, precision, scale, etc.

** Location Attributes such as device type, organization of storage, access method, addressing algorithm, activity address, block size, storage physical sequence, and directory aliases.

** Relationship Attributes including:

*** Logical Structure - pointer to subdivisions of subject record within a single system.

*** Membership - pointer to next highest sector/record of the same system under which the subject record is a member.

IV. CONTENTS AND PRODUCTS (Cont'd)

*** Interaction - pointer to a related item at equal hierarchical level in a different system.

** Organization Attributes such as expected occurrences, maximum occurrences, growth factor, frequency of use, overflow, priority, and statistics.

** Security Attributes include attributes relating to:

*** Security of the system

*** Security of the data content

*** Access authority

*** Data Source, Update, and Definition Responsibility

*** Privacy Considerations

*** Freedom of Information Considerations

** Integrity Attributes such as validity, consistency, reasonableness, propagation set, completeness, reliability, recovery and retention.

** Cost Attributes including such as development, design, production, overhead, maintenance, distribution, communications, retrieval, and support costs.

* Tables - A series of thirty-eight (tables) are incorporated in the data base which provide clear text translation of coded attributes and are used to edit inputs both as to accuracy and validity of select combinations of codes.

* Program Segments are also stored in the data base.

Data Relationships

* Physical records are stored indexed to an internal control number.

* Logical records consist of sets of physical records indexed to the same internal control number.

* Table records are likewise logical records consisting of three physical records, one for the data code, the second for the abbreviation and the third for the data value.

IV. CONTENTS AND PRODUCTS (Cont'd)

* By reason of the indexing of attributes internally, it is possible to query on a key and find all records related thereto. Virtually all attributes are key. Query is on a boolean and/or/not thus facilitating selective retrieval. Additionally, the system automatically develops descriptor (key word) indexing of the official and synonymous names thus offering increased potential for relating records. Further, the system develops invisible stemmed descriptors of selected key entries (e.g., first five positions of key word; first six positions of synonymous name key word; first four positions of system control designation; etc.). These further facilitate selective retrieval and avoid grammatical variances.

Report Products

Report products, at the present time are based upon ad-hoc queries. In this regard, it should be noted that the query feature of the system permits the user to define the search strategy, prescribe the sort key and delineate the print specification including entry of fixed labels. It is planned that program segments will be written to provide necessary preplanned reports.

V. OPERATIONAL CHARACTERISTICS

Data Collection

* All LOGDRMS updates are accomplished by direct on-line entry to the data bank with no intermediate handling.

* Add (create) actions are generally entered in two operations. The first creates a specified number of logical records with common attributes (parameter). The follow-on action is an update to add the unique (variables) attributes of each record. The program for additions prompts the field-name and solicits an entry or carriage return. A control matrix monitors the applicability of attributes to a particular segment.

* Changes are entered in one of two modes. Mass change of mulitiple records with identical information is accomplished as a single action. Updating of records requires identity of the fieldname to be changed and when ready the entry of the data applicable to the specified attribute.

* Edit/validation control using tables prevents entry of invalid codes or code combinations.

V. OPERATIONAL CHARACTERISTICS

* Changes to text such as Official Name, Synonymous Names, Definition, and Remarks are made with a text editing routine which allows for insertion, deletion, or replacement of only the changed characters without requiring any action on the unchanged portion of the text.

Operating Mode

* LOGDRMS is organized in three subsystems; (a) update; (b) query; and (c) batch processing products.

* All data updates (additions, changes, deletions) are performed on line direct to the data bank.

* Query is performed on-line interactive with search strategy defined in terms of fieldnames and key values utilizing boolean and; or; and not.

* Batch processing products are developed by extracting data base records and dumping same on a magnetic tape in response to a query command. The magnetic tapes are batch processed and prepared for entry into Computer Output Microfilm (COM) equipment which produces master microfiche which in turn are copied and distributed.

Interfaces

* At the present time, LOGDRMS does not directly interface with any other system.

* Future planning for enhancement calls for tape products from systems design centers to be batch loaded into LOGDRMS. Additionally, data terminals are to be installed at system design centers to make the LOGDRMS available on site of system development and to eliminate the need for hard copy (microfiche) publications for screening purposes.

VI. SECURITY

Restrictions on Use

* LOGDRMS is designed to document unclassified information.

* Restrictions are included by password control on update or query. Since sorting involves so much CPU time, separate passwords control who can sort retrieved data.

VI. SECURITY

Privacy Act Considerations

A complete section of a logical records is devoted to security attributes including privacy and freedom of information considerations.

Department of Defense - U.S. Navy

Navy World-Wide Military Command and Control System

Standard Data Element System

STADES

95

I. IDENTIFICATION

Office of Prime Responsibility

The Standard Data Element System (STADES) was developed by the Navy Regional Data Automation Center, Washington, DC (NARDAC, WASH., DC), (formerly the Naval Command Systems Support Activity (NAVCOSSACT)) to support the development, documentation, implementation, and control of the Navy World-Wide Military Command and Control System (WWCCS). For information on STADES contact:

Commanding Officer
Navy Regional Data Automation Center, Washington, DC
Attention: Code 70.3 (Mr. Frank Tagler)
Washington Navy Yard, Building 196
Washington, DC 20374

Telephones: AUTOVON: 288-3629
Commercial: (202) 433-3571

Documentation

System documentation is internally developed and maintained. The principal publication for the Navy WWMCCS Standard Data Element System (STADES) is NAVCOSSACT Document Number 88T001 TN-01 currently being revised as a NARDAC, Wash., DC publication.

II. SCOPE

Reason for Implementation

The World-wide Military Command and Control System is designed to promote commonality of computer programs and systems across command lines. An essential part of this effort is encouraging the use of existing data, data structures, and data files in order to simplify interchange of data and programs. The Navy WWMCCS Standard Data Element System (STADES) has been designed to facilitate this effort and to ease the problems of determining the structure of data and data codes that are used in programs and systems designed for WWMCCS use particularly when such data have been used in previously developed projects.

II. SCOPE (Cont'd)

Responsibility

The Navy Regional Data Automation Center, Wash., DC (NARDAC, Wash., DC), Washington Navy Yard, an element of the newly formed Naval Data Automation Command (NAVDAC), is responsible for management, maintenance, and promulgation of the Navy WWMCCS Standard Data Element System (STADES). As a part of this function, NARDAC, Wash., DC:

* Has established the initial centralized NAVY WWMCCS standard data element file from data elements in SECNAV Instruction 5200.20A, JCS Pub 6, JCS Pub 7, Federal Information Processing Standards (FIPS) and the DIA IDEAS file.

* Has developed and is maintaining procedures for the Navy WWMCCS STADES using RAS II to provide:

  ** Efficient and effective storage, indexing and data element classification for centralized NAVY WWMCCS data standards.

  ** Responsive inquiry and retrieval techniques for use by a local WWMCCS data manager.

* Determination of when local data element development should be authorized.

* Promulgation of RAS II and files, including updates, to Navy and Navy supported users.

Record Association System (RAS)

The Record Association System (RAS) II, the program system used for maintaining the WWMCCS STADES, is an automated cataloging and indexing system with extensive query and retrieval capabilities. The design of RAS II is based upon a capability which allows development of many separate data bases (modules) which are used independently or merged in selected combinations. In this manner, each user can control and maintain both EAM (Electrical Accounting Machine) and computer operations required for his use of the RAS II. However, the users modules are considered as mergeable files of an overall

II. SCOPE (Cont'd)

RAS II data base. Use of one set of record formats and programs in RAS II facilitates the exchange of information among all RAS II users. Individual access to the modules by any RAS II user is coordinated and accomplished by the NARDAC, Wash., DC RAS II staff. The NARDAC, Wash., DC RAS II staff also maintains two modules. One contains a file of approved standard data elements and related features and the other a file of information on all approved standards for ADP. These modules are available to all users for automatic merging with their respective data base modules. This capability eliminates redundant preparation of approved standards data by RAS II users.

When RAS II is used as a data element library system, initial investigations require that analysts collect information describing data systems or subsystems, the data files or reports which are part of the systems or subsystems and the data chains and data elements which constitute each file and report. Each data chain or data element is encoded so as to reference its related data chain standard or data element standard. Updates to a RAS II Data Base are facilitated by the automatic assignment of a permanent key (Record Identification Code) for identifying each record.

Three types of printed forms are used by the collectors. When the forms are completed they are submitted for review to a central coordinator familiar with both RAS II and the system. Upon completion of review, the forms are returned to the collectors for keypunching and verification. Sorting the cards prior to listing is optional but makes proofreading easier. The listing of the punched cards is reviewed by the collector and corrections or additions are provided as needed. The completed cards are submitted by the collector to the RAS II maintenance personnel for a generate or update computer run. The resulting listing shows the records modified by that run and any input errors encountered. Separate programs are used to format and list all or part of the data base. Catalog and retrieval runs are made when summary information from the data base is desired.

RAS II can also be employed as a tool for operations research efforts related to data requirements and data resources. Application of the RAS II methodology to define the elements of system data requirements and data resources provides an organized structure for analyses. Through the analysis of RAS II outputs redundant data reporting can be found and information requirements can be matched to existing data resources.

II. SCOPE (Cont'd)

RAS III, an updated version of RAS II, allows for online interactive query and updating of the RAS data base.

STADES Procedures

STADES is essentially a set of procedures governing the development and use of the data content of RAS II files with special emphasis on review of data elements and their related features during specified events of the systems development life cycle.

The Data Element Dictionary

The Data Element Dictionary is a structured output product of the system used to communicate information pertaining to standard date elements and related features to various classes of users. However, it should be noted that the RAS II capability allows for output of a wide range of information products other than the data element dictionary.

III. ADP RESOURCES

Although RAS II is available on the UNIVAC 1100 series, IBM 360/370, and CDC 6700 with slightly different capabilities, the system for Navy WWMCCS is run on the Honeywell 6000 series.

Hardware Requirements

* Honeywell 6000 series
* One (1) magnetic tape
* Three (3) tape drives
* One (1) card printer
* One (1) line printer
* One (1) data terminal
* Core memory: less than 29K
* Operating System: Honeywell GCOS

III. ADP RESOURCES (Cont'd)

Software Requirements

* All programs are written for ANSI COBOL

* Utilizes internally developed RAS II programs in three distinct blocks:

  ** Maintenance programs for generating and updating the data base.

  ** Catalog programs for formating and printing the data base.

  ** Retrieval programs which provide various query capabilities

IV. CONTENTS AND PRODUCTS

Purpose

RAS II has been implemented on Navy WWMCCS computers in support of STADES to provide:

* An automated inventory of NAVY WWMCCS data element information

* Indexes of this information

* Catalog Report Generation Capability

* Query/Retrieval Capability

Data Relationships

RAS is an automated data storage indexing and retrieval system. Its primary use in the Navy is to maintain data element libraries by generating files of data elements and related information, indexing the data elements, generating data element catalogs and retrieving data from the data element files. These files are modular, allowing separate organizations to independently develop their own data element libraries and yet be able to share data or files, by automated means, with other RAS data element libraries. These interacting data element libraries

IV. CONTENTS AND PRODUCTS (Cont'd)

make up a Navy-wide data element library system. The data stored in data element files are retrieved for specific analysis, design, and data standardization purposes. The data is indexed by assigned classifications and by keywords. Logical associations of these indexes complement a query capability for retrieval of the data. Catalogs, reference indexes, and specific retrievals of data may be generated. RAS is sufficiently flexible for storing, indexing, and retrieving data on entities other than data elements and related features. It is used for document cataloging and report control.

Data Organization

The WWMCCS STADES contains information about:

* Data Systems and Subsystems
* Data Files
* Data Record Types
* Data Use Identifiers and Data Chains
* Standard Data Elements and Standard Data Chains

RAS Programs

* Four maintenance programs
* Two maintenance programs for Distributive Data Base
* One associating program providing five reference indexes
* Six catalog report programs
* Five basic retrieval programs
* Two interactive programs

V. OPERATING MODE

Data Collection

* Creating an initial record is performed on a batch basis.

V. OPERATING MODE (Cont'd)

* Changes to existing records are processed on a batch basis (RAS II) and/or on line in the case of facilities utilizing the RAS III Interactive version.

* Controls are built in to edit updates and to control inputs only from those sources authorized to update the records involved.

Operating Mode

* RAS II is designed for batch processing (update, query, and display).

* RAS III provides on line interactive query capability as well as on line change capability.

Interfaces

The WWMCCS STADES is designed to interface with user systems (e.g., WWMCCS, CINCPAC, COMSUBLANT, DCA JTSA, etc.).

VI. SECURITY

Restrictions on Use

* STADES includes provisions for recording the security classification of the data or entity being described.

* STADES also includes controls to assure that only authorized activities can introduce or update records.

Privacy Act Considerations

Considerations of privacy can be incorporated in remarks notes of applicable records.

Department of Agriculture

Data Base Directory

USDADBD

USDA Data Base Directory

I. IDENTIFICATION

The USDA Data Base Directory is being developed by the USDA Office of Automated Data Systems (ADS). The directory was first designed in 1975 to provide an inventory of on-line applications operated by the various independent agency data processing staffs. The development of this inventory was conducted by the:

Plans and Policy Division
Office of Automated Data Systems
U. S. Department of Agriculture
Washington, D. C. 20250

Roxanne Williams, Director of Plans and Policy Division, ADS, USDA is the contact point for information concerning the Data Base Directory (202/447-2118). Copies of the data base design used to support this inventory can be obtained by contacting ADS directly.

II. SCOPE

In 1972 the Department conducted a detailed survey of program data requirements. Both manual and automated data were identified according to the Departmental programs which they support. The following data attributes were collected, placed on-line on a 370/145 for retrieval, and published for information interchange in the USDA Data Inventory: title of program data requirement, descriptive abstract, major data elements, source of data and processing mode. This initial inventory was conducted to identify potential common interest data; thus, supporting the Department's move into the data base (DBMS) environment. Now that the Department has transaction processing and data base management capabilities, it has become more important to identify common data use and to catalog on-line data to increase information exchange.

The USDA Data Base Directory (DBD) is one tool which has been developed to augment the data management needs of the Departmental/ Agency Data Base Administration (DBD) staffs. The DBD was created as a Data Resource Management System or tool which can provide DBA staffs with information about on-line applications needed to manage data bases.

The objectives and operational uses for the DBD are to provide:

- Individual agency access to information about existing automated data in support of data interchange within the Department.
- Computer centers with software utilization figures.
- The Departmental DBA staff with a global view of processing trends.
- The DBA staff with an inventory of data bases Department-wide.

The DBD was developed in 1975 as a follow-up to the 1972 USDA Data Inventory study. The DBD contains information on approximately thirty-five data bases to date; the inventory has just been started and currently reflects only a small percentage of the Departmental data bases. The present structure of the DBD is under review by the agency data base managers and based on experience to date, is subject to modification.

III. ADP RESOURCES

The USDA DBD is operational on an IBM 370/168 using batch and on-line processing through TSO into a single user version of the Data Base Management System (System 2000 of MRI Systems Corporation).

IV. CONTENTS AND PRODUCTS

The data relationships utilized by the USDA DBD include:

- A repeating group concept using System 2000 that ties each data base or group of data bases to the system which it supports--showing the title and description of the system and the title and description of the supporting data bases.
- The major data elements contained in each data base.
- Titles of reports generated from the data bases.
- List of USDA agencies other than the owner agency interested in the data bases.

V. OPERATIONAL CONSIDERATIONS

Data collection for the DBD is accomplished using two methods. First, USDA procedures for development of on-line data bases require submission of data base specification to the Departmental DBA staff. These specifications and documentation information are entered into the DBD. Second, for data bases already in existence, agency DBA staffs collect the information and forward the data to the Plans and Policy Division of ADS for inclusion in the DBD.

VI. SECURITY

This data is then entered via TTY compatible terminal into the System 2000 data base. Control of the DBD access is accomplished using the System 2000 password security.

VII. USER CHARACTERISTICS

The Office of Automated Data Systems is a staff office with total responsibility for management of the Department's Data Processing operations. ADS works directly with the ADP staffs of the Departmental line agencies. ADS maintains the DBD, each agency has copies of the DBD output. At present, the DBD contains information about online data bases covering:

- Name of organization responsible for the data base system
- Agency contact
- System name
- Data base name
- Processing location
- Software used
- Data base size
- Data base description
- Major data elements
- Source of data elements
- Data use (which program it supports)
- Data base report
- Update mode and frequency
- Other interested USDA agencies
- Personnel data indicator
- Security features

This list is not all inclusive, and is subject to modification as the Department gains more experience in its use of the DBD.

VII. COSTS

Any analysis of USDA costs would be premature at this time since the existing data base is so small. The existing DBMS, System 2000, was adequate for our needs so that no special software acquistion costs were incurred.

Veterans Administration Data Dictionary

Veterans Administration

VADD

VETERANS ADMINISTRATION
VA DATA DICTIONARY

I. IDENTIFICATION

The Veterans Administration Data Dictionary (VADD), designed and developed by the Department of Data Management of the Veterans Administration, was installed July 15, 1976. This system is the responsibility of the Privacy and Data Administration Division (334), and for information, please contact:

Chief, Privacy and Data Administration Division
810 Vermont Avenue, N.W.
Washington, D. C. 20420

Phone: (202) 389-3034

Documentation, in the form of a "Handbook to the VA Data Dictionary", is available on request. This document is primarily a user's guide to the system.

II. SCOPE

The objectives of the VA Data Dictionary are:

- To standardize data elements through a data inventory, by first identifying and then isolating those data elements with common definitions and attributes.
- To identify and eliminate and/or prevent redundancy and inconsistencies in data elements used within VA ADP management.
- To provide detailed documentation of data elements, to insure that all pertinent facts about the elements are communicated directly and concisely to both ADP and non-ADP management.
- To aid VA personnel in determining the impact of proposed and/or approved changes to data elements by identifying the programs, master files and reports in which the data elements appear.
- To support VA personnel engaged in designing, maintaining, and managing automated systems by providing easily comprehensible reports describing the VA data resource.

The VA Data Dictionary is an information tool by which the Veterans Administration intends to manage and control the data resource. This information tool is an automated, central collection point for a current and complete description of all VA data elements. In addition to this dictionary feature, the system is also a data directory in that data elements are linked to application systems, user programs, master files and reports.

The total number of data elements in the VA is approximately 82,000 elements. As of November, 1976, there exist 374 application systems in the VA. This amount can easily be handled by the VA Data Dictionary. A time span of approximately two years is anticipated for entering all VA ADP applications into the VA Data Dictionary.

III. OPERATING REQUIREMENTS

The VA Data Dictionary is installed on an IBM Series 360/65 computer, running under OS/MVT. The system requires 90K bytes of main storage, a single 2314 disk (to be used for intermediate work files), two tape drives (three may be used as an alternative to disk for purposes of spooling output) and a printer. Optionally, the VA uses a Xerox 1200 off-line printing system.

The dictionary programs are written in ANS COBOL.

The software supporting the VA Data Dictionary is non-proprietary.

IV. CONTENTS AND PRODUCTS

The VA Data Dictionary Master File, stored on magnetic tape, is organized by sequence number within transaction code, within Permuted Word (key identifier for a data element), within application acronym. The record length is 125 positions, blocked 8. There are seven multiple fixed length master records that contain: specific information on a data elements (such as name, description, values, users, application programs affected); element content of an application's master files; and data elements that are used in the creation of application system reports. A Summary record follows each Application System storing necessary statistical data.

The system is designed to run in a batch mode with sequential update and report processing to occur a minimum of twice a week. It should be noted that frequent update of the master file will provide users with the most current, up-to-date information on data elements within their application(s).

Semi-annually the entire VA Data Dictionary is printed.

The VA Data Dictionary generates several reports on the data element contents of VA applications. Three distinct report types are produced: "Element Description Report", Master File Report", and "Report Content" report. These reports are generated bi-weekly as changes occur to the information content of a data element, master file, or report, and semi-annually when the entire contents of the Data Dictionary is printed.

The primary report generated by the VA Data Dictionary is an Element Description Report which describes each data element within a VA application. The information provided includes: Permuted Word (key identifier of the data element); Element Name; Source Responsibility (organizational element that controls or manages a VA Program Application); Narrative Description of the Element and its related values; Application Programs and VA departments that use the Data Element.

As stated previously, the VA Data Dictionary links data elements to application master files and reports. The "Master File Description Report: identifies, by application system, all data elements stored within a master file in its proper sequence. An "Element Structure Report: identifies, for each report generated by an application system, all data elements used in the creation of the report.

A by-product of this dictionary is the "permuted Word" Report. The "Permuted Word" Report lists all VA data elements, in Permuted Word sequence, and the application system that uses the data element. For a brief explanation, this provides a means of cross-referencing data elements common to multiple applications.

The update program generates separate lists, by application systems, for all accepted or rejected transactions. In addition, a statistical summary report lists, by application system, individual totals for the number of elements, master files and reports contained in the VA Data Dictionary Master File for an application, in addition to other miscellaneous statistical data.

The VA Data Dictionary provides a query facility through the use of optional reports. These reports, run on an IBM Series 360 computer or other compatible computer hardware, consist of specific application requests or partial reports about: (1) a specific data element or (2) the contents of a specific application master file, or (3) the data elements used in the construction of each report. Storage of the Master Files on tapes makes it very conducive to distribution of these files to the various VA Data Processing Centers throughout the United States. This will permit each Data Processing Center to obtain any portion of the Dictionary that is in current demand.

The system does not have an RPG capability nor provide any interactive activity, special utilities or interfaces with any other system. However, at a later time, certain information may be extracted from the system and used as input to a Data Base Management System.

V. OPERATIONAL CONSIDERATIONS

The organization responsible for management and control over a VA application (the "Responsible Source") will designate certain personnel to prepare input data to the VA Data Dictionary. These people, commonly referred to as "Information Specialists", should be familiar with and knowledgeable of the information contents of their application systems.

Data to be entered into the VA Data Dictionary System initially is obtained from existing system documentation found in manuals, circulars and directives, or through the expertise of various individuals. Appropriate systems personnel at a VA Data Processing Center (DPC) may provide data such as application program names, master files and reports content. This data gathering and preparation can be phased in at different time periods. That is, data element name and description will establish a data element on the master file with additional information to be provided at a later time.

Input to the Data Dictionary system is accomplished through key-entry equipment using seven different transaction formats.

Procedures have been established to insure integrity of data to be submitted to an processed by the VA Data Dictionary. These include various circulars and the Data Dictionary Handbook. Continued usage of the dictionary by the responsible sources of information will enhance the integrity of the dictionary.

The VA Data Dictionary edits not only for validity, but completeness of data submitted for processing. Any free-formatted areas are not edited. Back-up to the master file is provided through retention of father and grandfather master files and transaction tapes.

VI. SECURITY

Procedures have been established for the development of an access to the VA Data Dictionary. Access is available to all authorized personnel in the Veterans Administration. Tape copies of the master file are available to all DPC's for query and optional reports at a user's request. Since the dictionary is a "hard-copy" version, limitation of access is not intended to be restrictive; since at the present time, it is not intended to be a data base system. A provision has been made to indicate the security level of each data element. At this time, however, it is not an active feature in the element description.

All input documents are submitted through the Privacy and Data Administration Division for review, analysis and control. Computer generated reports are distributed to authorized personnel according to established procedures.

VII. USERS

The VA Data Dictionary is a useful information tool, to be used VA-wide and not limited to a specific application or geographical area. Because of the layman's language description and definition of data elements, the dictionary may be utilized by ADP oriented as well as non-ADP oriented management and staff. As a result, the dictionary may be used as a reference point for communication among management, "Responsible Source", users, programmers and analysts. It also will be used as the basic ingredient in formalizing a data base system.

APPENDIX D
SCOPE NOTES FOR THE
NARRATIVE SYSTEM DESCRIPTIONS

I. IDENTIFICATION

A. Title. Full system title.

B. Short Title. Commonly used abbreviation or acronym identifying the system.

C. Organization and Address. Identity of the organization responsible for the development of the system to the Branch or Division level and complete mailing address.

D. Technical Contact Title, Organization, and Phone Number. Identity of the office to be contacted for detailed information on the subject matter or operational aspects of the system. Include area code and phone number and phone system (i.e., FTS, AUTOVON, commercial).

E. Documentation Type and Availability. Briefly describe available documentation. Classify the type of document (i.e., functional description, system design specification, user manual, etc.) document title, document reference number, and indicate the availability of documentation listed. Identify the contact point for obtaining documentation.

II. SCOPE

A. Background. Briefly describe the circumstances which resulted in the development of a DED/D system. If it is not the organization's first attempt at organized data management, describe the previous efforts:

1. History of DED/D system if obtained from another organization or modified an already existing system.

2. Approximate month and year in which major system milestones occurred or are expected to occur.

3. Reasons for adoption, creation, or modification of DED systems.

B. Implementation Level. Determine the level of capability as listed and defined in Section 2 of this report. The category selected should reflect accurately the system's primary function in the view of the developer.

III. ADP RESOURCES (IF NEEDED)

A. HARDWARE RESOURCES

Describe the hardware and software requirements of automated systems. The hardware resource section should include the manufacturer's name and model number of all computer main frames on which, to the knowledge of the developer, the system is installed.

The operating system used and the minimum core required to process the system should be specified. Both standard peripherals, such as disk and tape drives, and special offline devices, such as page printers or microform devices, which are needed to operate the system may be listed.

B. Software

Describe the software resources in two distinct groups. The first group should include any programs written by or extensively modified by the organization which installed the system such as edit and update programs and report generation programs.

The second group should include any generalized software needed to execute programs in the first group. These would include data base management systems, generalized report generator modules, program library facilities, and utilities such as sorts and data conversion routines.

For each group the developer's name, the source language, its availability and its proprietary status and cost, if applicable, should be listed. If a group of proprietary programs is involved it is listed as a group, not by component programs. Only user modifications should be mentioned.

IV. CONTENTS AND PRODUCTS

A. Data Relationship Described

In this section, all entities (as previously defined by task group 17) such as reports files, programs etc., which are covered by the system, can be listed. The relationships between the information entities which the DED/D describes may also be outlined. For example, if data elements are related to reports, or files to programs, this may be recorded.

B. Data Organization

Summarize the content and organization of the DED/D and give a brief description of how it is physically stored and organized, and whether or not it is stored in machine readable form.

List all files, their primary purpose, and their content (record and field) with sequence and access methods. For example, a system might have one file which describes data elements and another which describes reports. For each file list the fields for each record in the file such as element name, size, or class, or report names, report data etc, and indicate sequence fields. Specify the storage medium used for the file, i.e., microfiche, paper, tape, disk, etc.

C. Reports and Products

List all standard products produced by the system and briefly describe them if the title is not self-explanatory. Outline any ad hoc reporting and query capability.

V. OPERATIONAL CONSIDERATIONS

A. Data Collection

Cover the sources of the data used to build and maintain the DED/D such as data processing or forms management, data processing users, systems design personnel, data base administration staff, etc.

Discuss the methods used to collect data. Also describe data collection methodologies, such as research through system documentation, interviews conducted by analysts, or the use of special forms designed for the purpose. Describe the skill level and the procedures used in the entry and update processes.

B. Operating Mode

Mention the system's operating mode, both update and query.

Is it batch or online for both or is it a mixture? Briefly outline this capability.

C. Interfaces

List interfaces with other application systems or generalized software. An example of this would be an interface between the DED/D and a financial management system or a project control system.

VI. SECURITY

A. Restriction on Use

Discuss any restrictions on use of the DED/D, the reasons for the restrictions, and the ways they are enforced. List the Personnel and offices who normally perform updates.

B. Privacy Act of 1974 Considerations

Cover any effects which the organization feels that the Privacy Act of 1974 has had upon the way in which the DED is used or constructed.

VII. User Characteristics

A. User Organization. Describe of the organizational relationship of the responsible organization to user organizations, and the level of the responsible organization within the overall organization.

B. Data Used/Maintained. List the information entered and maintained by the DED/D for user organizations.

C. Reasons for Use. List the reasons for using the DED/D by user organizations.

VIII. COSTS (OPTIONAL)

A. Cost. Supply any available estimates of cost in dollars which may be helpful to prospective users or developers in scoping resource commitments if they expect to develop and maintain approximately the same number of data elements and expect similar amounts of activity against the DED/D.

Include costs for equipment, software, maintenance, and related support costs.

B. Personnel Committment.

1. List the number and type or classification of personnel used during each phase of system development and operations.

2. List skill levels by type or classification of personnel employed during development and operation.

3. List of tasks performed during development and operation.

IX. General Comments

Include additional comments which the developer wishes to make.

APPENDIX E
DEFINITIONS OF INFORMATION ENTITIES

The following definitions are intended to promote a clear understanding of the meaning of each one of the Entity Classes that are identified in this publication. It should be understood that what constitutes an Entity Class is environment dependent. What one organization may determine as appropriate classes of entities and their relationship to one another may differ from what another organization considers appropriate.

Plan/Program - An activity directed toward a common purpose, objective, or goal undertaken or proposed by an organization in order to carry out responsibilities assigned to it. In many instances the plan or program has its root in legislation or executive determination. (See Plan [Management] and Program [Management].)

Plan (Management) - A documentation of management goals together with a general prescription of the manner in which such goals will be attained within a defined timeframe.

Program (Management) - A combination of activities undertaken by one or more organizations directed toward attaining specified management objectives in given functional and subject matter areas. Initiation of a management program is normally effected by issue of a directive, instruction, regulation, or other type of authoritative issuance which prescribes the scope of efforts, enunciates management objectives and program policies, assigns responsibilities, and authorizes the use of resources for the purposes of the program.

System - A composite of equipment, skills, techniques, and information capable of performing and/or supporting an operation role in attaining specified management objectives. A complete system includes related facilities, equipment, material, services, personnel, and data/information required for its operation to the degree that it can be considered a self-sufficient unit in its intended operational and/or support environment.

Application - The first level subdivision of a system consisting of a series of processes or procedures devoted to accomplishing a specified portion but not all of the system objectives. In management systems, applications are often identified as subsystems. In automated data processing systems, applications are frequently identified as either subsystems or application systems.

Procedure - A series of precise step-by-step processes within an application which produce specified results. The processes may be manual or automated or a combination of both. The terminology normally used to describe an automated process is "computer program"; manual processes are "tasks".

Process (Operational) - The first level subdivision of an application consisting of a systematic sequence of operations to produce a specified result. (See also Procedure).

File - A collection of related records which are treated as a unit. (With the advent of computers, it has now become feasible to collect related records of one or more files to be treated as a unit to serve one or more processes. This form of related file structure is called a DATA BASE. In the event a decision is made to record data bases as a separate entity class in a DRD, they should be recorded at a hierarchically higher level than FILES.)

Record - A collection of related data elements treated as a unit.

Record (File) - A collection of related file items of data treated as a unit.

Data Element - A basic unit of identifiable and definable information. A data element occupies specific space on a form, report, or record. It has an identifying name and value or values for expressing a specific fact.

Report - A product of a process which provides, a narrative, statistical, or graphical presentation of one or more records and/or other information transmitted for use in planning, controlling or evaluating operations and performance, and determining policy.

Form - Any printed design with or without text which contains blank spaces to be filled in to record, collect, or transmit one or more records. A form completed by the entry of information may be a file record, a document, or a report.

Document (Transaction) - A collection of related items of data treated as a unit for input to a system, application, or process.

NBS-114A (REV. 7-73)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET	1. PUBLICATION OR REPORT NO. NBS SP-500-16	2. Gov't Accession No.	3. Recipient's Accession No.

4. TITLE AND SUBTITLE COMPUTER SCIENCE & TECHNOLOGY:
A Survey of Eleven Government-Developed Data Element Dictionary/Directory Systems

5. Publication Date: August 1977

6. Performing Organization Code

7. AUTHOR(S) Hazel McEwen, Coordinator for FIPS Task Group 17C - Data Element Dictionary/Directory

8. Performing Organ. Report No.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

10. Project/Task/Work Unit No. 640.1131

11. Contract/Grant No.

12. Sponsoring Organization Name and Complete Address (*Street, City, State, ZIP*)
Same as No. 9

13. Type of Report & Period Covered: Final

14. Sponsoring Agency Code

15. SUPPLEMENTARY NOTES
Library of Congress Catalog Card Number: 77-608155

16. ABSTRACT (*A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.*)

This report presents the current state-of-the-art of government-developed Data Element Dictionary/Directory (DED/D) systems. DED/D's are software tools used for managing and controlling information and data. Eleven DED/D systems are described, first using a side-by-side features presentation approach, and followed by narrative systems descriptions which highlight special capabilities and experiences with each system. Information presented in this report is intended to serve both the technical and administrative ADP community.

17. KEY WORDS (*six to twelve entries; alphabetical order; capitalize only the first letter of the first key word unless a proper name; separated by semicolons*)
ADP; automated data processing; computer software; data element dictionary; data element dictionary/directory; data management; software tool.

18. AVAILABILITY
[X] Unlimited
[ ] For Official Distribution. Do Not Release to NTIS
[X] Order From Sup. of Doc., U.S. Government Printing Office Washington, D.C. 20402, SD Cat. No. C13.10:500-16
[ ] Order From National Technical Information Service (NTIS) Springfield, Virginia 22151

19. SECURITY CLASS (THIS REPORT)	21. NO. OF PAGES
UNCLASSIFIED	111
20. SECURITY CLASS (THIS PAGE)	22. Price
UNCLASSIFIED	

USCOMM-DC 29042-P74

U. S. GOVERNMENT PRINTING OFFICE : 1977--240-848/266